1461599

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

May 03, 2006

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: *60/503,460*
FILING DATE: *September 17, 2003*
RELATED PCT APPLICATION NUMBER: *PCT/US04/30397*

THE COUNTRY CODE AND NUMBER OF YOUR PRIORITY APPLICATION, TO BE USED FOR FILING ABROAD UNDER THE PARIS CONVENTION, IS *US60/503,460*

Certified by

Jon W. Dudas

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

Mail Stop Provisional Patent Application

PTO/SB/16 (6-95)
Approved for use through 04/11/98. OMB 0651-0037
Patent and Trademark Office, U.S. DEPARTMENT OF COMMERCE

PROVISIONAL APPLICATION COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION under 37 CFR 1.53 (c).

	Docket Number	1579-854	Type a plus sign (+) inside this box→	+

INVENTOR(S)/APPLICANT(S)

LAST NAME	FIRST NAME	MIDDLE INITIAL	RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY)
Client File No. 1579854			

TITLE OF THE INVENTION (280 characters)

CONSENSUS/ANCESTRAL IMMUNOGENS

CORRESPONDENCE ADDRESS

Direct all correspondence to:

☒ **Customer Number:** 23117

Type Customer Number here

Place Customer Number Bar Label Here →

ENCLOSED APPLICATION PARTS *(check all that apply)*

☒ Specification	Number of Pages	67	☐ Applicant claims "small entity" status. ☐ "Small entity" statement attached.	
☒ Drawing(s)	Number of Sheets	50	☒ Other (specify)	ABSTRACT

METHOD OF PAYMENT *(check one)*

☒ A check or money order is enclosed to cover the Provisional filing fees ($160.00)/($80.00)	PROVISIONAL FILING FEE AMOUNT ($)	160.00
☐ The commissioner is hereby authorized to charge filing fees and credit Deposit Account Number 14-1140		

The invention was made by an agency of the United States Government or under a contact with an agency of the United States Government.

☐ No.

☒ Yes, the name of the U.S. Government agency and the Government contract number are:

National Institutes of Health, Contract No. NO1AI 85338

Respectfully submitted,

SIGNATURE Mary J. Wilson DATE September 17, 2003

REGISTRATION NO. *(if appropriate)* 32,955

TYPED or PRINTED NAME Mary J. Wilson

☐ Additional inventors are being named on separately numbered sheets attached hereto.

PROVISIONAL APPLICATION FILING ONLY

Burden Hour Statement: This form is estimated to take .2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Mail Stop Comments - Patents, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, and to the Office of Information and Regulatory Affairs, Office of Management and Budget (Project 0651-0037), Washington, DC 20503. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Mail Stop Provisional Patent Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

775847

Atty. Dkt. 1579-854

U.S. PROVISIONAL PATENT APPLICATION

Inventor(s): Client File No. 1579854

Invention: CONSENSUS/ANCESTRAL IMMUNOGENS

NIXON & VANDERHYE P.C.
ATTORNEYS AT LAW
1100 NORTH GLEBE ROAD, 8TH FLOOR
ARLINGTON, VIRGINIA 22201-4714
(703) 816-4000
Facsimile (703) 816-4100

SPECIFICATION

CONSENSUS/ANCESTRAL IMMUNOGENS

TECHNICAL FIELD

The present invention relates, in general, to an immunogen and, in particular, to an immunogen for inducing antibodies that neutralize a wide spectrum of HIV primary isolates and/or to an immunogen that induces a T cell immune response. The invention also relates to a method of inducing anti-HIV antibodies, and/or to a method of inducing a T cell immune response, using such an immunogen. The invention further relates to nucleic acid sequences encoding the present immunogens.

BACKGROUND

The high level of genetic variability of HIV-1 has presented a major hurdle for AIDS vaccine development. Genetic differences among HIV-1 groups M, N, and O are extensive, ranging from 30% to 50% in *gag* and *env* genes, respectively (Gurtler et al, J. Virol. 68:1581-1585 (1994), Vanden Haesevelde et al, J. Virol. 68:1586-1596 (1994), Simon et al, Nat. Med. 4:1032-1037 (1998), Kuiken et al, Human retroviruses and AIDS 2000: a compilation and analysis of nucleic acid and amino acid sequences (Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico)). Viruses within group M are further classified into nine genetically distinct subtypes (A-D, F-H, J and K) (Kuiken et al, Human

retroviruses and AIDS 2000: a compilation and analysis of nucleic acid and amino acid sequences (Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico, Robertson et al, Science 288:55-56 (2000), Robertson et al, Human retroviruses and AIDS 1999: a compilation and analysis of nucleic acid and amino acid sequences, eds. Kuiken et al (Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico), pp. 492-505 (2000)). With the genetic variation as high as 30% in *env* genes among HIV-1 subtypes, it has been difficult to consistently elicit cross-subtype T and B cell immune responses against all HIV-1 subtypes. HIV-1 also frequently recombines among different subtypes to create circulating recombinant forms (CRFs) (Robertson et al, Science 288:55-56 (2000), Robertson et al, Human retroviruses and AIDS 1999: a compilation and analysis of nucleic acid and amino acid sequences, eds. Kuiken et al (Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico), pp. 492-505 (2000), Carr et al, Human retroviruses and AIDS 1998: a compilation and analysis of nucleic acid and amino acid sequences, eds. Korber et al (Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico), pp. III-10-III-19 (1998)). Over 20% of HIV-1 isolates are recombinant in geographic areas where multiple subtypes are common (Robertson et al, Nature 374:124-126 (1995), Cornelissen et al, J. virol.

70:8209-8212 (1996), Dowling et al, AIDS 16:1809-1820 (2002)), and high prevalence rates of recombinant viruses may further complicate the design of experimental HIV-1 immunogens.

To overcome these challenges in AIDS vaccine development, the use of centralized HIV-1 genes (consensus, ancestral and center of the tree) to decrease genetic distances between candidate immunogens and field viral strains has been proposed (Gaschen et al, Science 296:2354-2360 (2002), Gao et al, Science 299:1517-1518 (2003), Nickle et al, Science 299:1515-1517 (2003)). Centralized sequences are located close to the root of HIV-1 phylogenetic trees and can significantly decrease genetic distances to contemporary field HIV-1 isolates.

The present invention results, at least in part, from the results of studies designed to determine if centralized immnunogens can induce both T and B cell immune responses in animals. These studies have involved the generation of an artificial group M consensus *env* gene (CON6), and construction of DNA plasmids and recombinant vaccinia viruses to express CON6 envelopes as soluble gp120 and gp140CF proteins. The results demonstrate that CON6 Env proteins are biologically functional, possess linear, conformational and glycan-dependent epitopes of wild-type HIV-1, and induce cytokine-producing T cells that recognize T cell epitopes of both HIV subtypes B and C. Importantly, CON6 gp120 and gp140CF proteins induce

antibodies that neutralize subsets of subtype B and C HIV-1 primary isolates.

SUMMARY OF THE INVENTION

The present invention relates to an immunogen for inducing antibodies that neutralize a wide spectrum of HIV primary isolates and/or to an immunogen that induces a T cell immune response, and to nucleic acid sequences encoding same. The invention also relates to a method of inducing anti-HIV antibodies, and/or to a method of inducing a T cell immune response, using such an immunogen.

Objects and advantages of the present invention will be clear from the description that follows.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-1D: Generation and expression of the group M consensus *env* gene (CON6). The complete amino acid sequence of CON6 gp160 is shown. (Fig. 1A) The five regions from the wild-type CRF08_BC (98CN006) *env* gene are indicated by underlined letters. Variable regions are indicated by brackets above the sequences. Potential N-liked glycosylation sites are highlighted with bold-faced letters. (Fig. 1B) Constructs of CON6 gp120 and gp140CF. CON6 gp120 and gp140CF plasmids were engineered by introducing a stop codon after the gp120 cleavage site or before the transmembrane domain, respectively. The gp120/gp41 cleavage site and fusion domain of gp41 were deleted in the

gp140CF protein. (Fig.1C) Expression of CON6 gp120 and gp140CF. CON6 gp120 and gp140CF were purified from the cell culture supernatants of rVV-infected 293T cells with *galanthus Nivalis argarose* lectin columns. Both gp120 and gp140CF were separated on a 10% SDS-polyarylamide gel and stained with Commassie blue. (Fig. 1D.) CON6 *env* gene optimized based on codon usage for highly expressed human genes.

Figures 2A-2E. Binding of CON6 gp120 gp140 CF to soluble CD4 (sCD4) and anti-Env mAbs. (Figs. 2A-2B) Each of the indicated mabs and sCD4 was covalently immobilized to a CM5 sensor chip (BIAcore) and CON6 gp120 (Fig. 2A) or gp140CF (Fig. 2B) (100 μg/ml and 300 μg/ml, respectively) were injected over each surface. Both gp120 and gp140CF proteins reacted with each anti-gp120 mabs tested except for17b mab, which showed negligible binding to both CON6 gp120 and gp140CF. To determine induction of 17b mab binding to CON6 gp120 and gp140CF, CON6 gp120 (Fig. 2C) or gp140CF (Fig. 2D) proteins were captured (400-580 RU) on individual flow cells immobilized with sCD4 or mabs A32 or T8. Following stabilization of each of the surface, mAb 17b was injected and flowed over each of the immobilized flow cells. Overlay of curves show that the binding of mab 17b to CON6 Env proteins was markedly enhanced on both sCD4 and mab A32 surfaces but not on the T8 surface (Figs. 2C-2D). To determine binding of CON6 gp120 and gp140CF to human

mabs in ELISA, stock solutions of 20μg/ml of mabs 447, F39F, A32, IgG1b12 and 2F5 on CON6 gp120 and gp140CF were tittered (Fig. 2E). Mabs 447 (V3), F39F (V3) A32 (gp120) and IgG1b12 (CD4 binding site) each bound to both CON6 gp120 and 140 well, while 2F5 (anti-gp41 ELDKWAS) only bound gp140CF. The concentration at endpoint titer on gp120 for mab 447 and F39F binding was <0.003 μg/ml and 0.006 μg/ml, respectively; for mab A32 was <0.125 μg/ml; for IgG1b12 was <0.002 μg/ml; and for 2F5 was 0.016 μg/ml.

Figures 3A and 3B. Infectivity and coreceptor usage of CON6 envelope. (Fig. 3A) CON6 and control *env* plasmids were cotransfected with HIV-1/SG3Δenv backbone into human 293T cells to generate Env-pseudovirions. Equal amounts of each pseudovirion (5 ng p24) were used to infect JC53-BL cells. The infectivity was determined by counting the number of blue cells (infectious units, IU) per microgram of p24 of pseudovirons (IU/μg p24) after staining the infected cells for β-gal expression. (Fig. 3B) Coreceptor usage of the CON6 *env* gene was determined on JC53BL cells treated with AMD3100 and/or TAK-799 for 1 hr (37°C) then infected with equal amounts of p24 (5 ng) of each Env-pseudovirion. Infectivity in the control group (no blocking agent) was set as 100%. Blocking efficiency was expressed as the percentage of IU from blocking experiments compared

to those from control cultures without blocking agents. Data shown are mean ± SD.

Figure 4. Western blot analysis of multiple subtype Env proteins against multiple subtype antisera. Equal amount of Env proteins (100 ng) were separated on 10% SDS-polyacrylamide gels. Following electrophoresis, proteins were transferred to Hybond ECL nitrocellulose membranes and reacted with sera from HIV-1 infected patients (1:1,000) or guinea pigs immunized with CON6 gp120 DNA prime, rVV boost (1:1,000). Protein-bound antibody was probed with fluorescent-labeled secondary antibodies and the images scanned and recorded on an infrared imager Odyssey (Li-Cor, Lincoln, NE). Subtypes are indicated by single-letters after Env protein and serum IDs. Four to six sera were tested for each subtype, and reaction patterns were similar among all sera from the same subtype. One representative result for each subtype serum is shown.

Figure 5. T cell immune responses induced by CON6 Env immunogens in mice. Splenocytes were isolated from individual immunized mice (5 mice/group). After splenocytes were stimulated *in vitro* with overlapping Env peptide pools of CON6 (black column), subtype B (hatched column), subtype C (white column), and medium (no peptide; gray column), INF-γ producing cells were determined by the ELISPOT assay. T cell IFN-γ responses induced

by either CON6 gp120 or gp140CF were compared to those induced by subtype specific Env immunogens (JRFL and 96ZM651). Total responses for each envelope peptide pool are expressed as SFCs per million splenocytes. The values for each column are the mean ± SEM of IFN-γ SFCs (n=5 mice/group).

Figures 6A-6E. Construction of codon usage optimized subtype C ancestral and consensus envelope genes (Figs. 6A and 6B, respectively). Ancestral and consensus amino acid sequences (Figs. 6C and 6D, respectively) were transcribed to mirror the codon usage of highly expressed human genes. Paired oligonucleotides (80-mers) overlapping by 20 bp were designed to contain 5' invariant sequences including the restriction enzyme sites EcoRI, BbsI, Bam HI and BsmBI. BbsI and BsmBI are Type II restriction enzymes that cleave outside of their recognition sequences. Paired oligomers were linked individually using PCR and primers complimentary to the 18 bp invariant sequences in a stepwise fashion, yielding 140bp PCR products. These were subcloned into pGEM-T and sequenced to confirm the absence of inadvertant mutations/deletions. Four individual pGEM-T subclones containing the proper inserts were digested and ligated together into pcDNA3.1. Multi-fragment ligations occurred repeatly amongst groups of fragments in a stepwise manner from the 5' to the 3' end of the gene until the entire gene was

reconstructed in pcDNA3.1. (See schematic in Fig. 6E.)

Figure 7. JC53-BL cells are a derivative of HeLa cells that express high levels of CD4 and the HIV-1 coreceptors CCR5 and CXCR4. They also contain the reporter cassettes of luciferase and β-galactosidase that are each expressed from an HIV-1 LTR. Expression of the reporter genes is dependent on production of HIV-1 Tat. Briefly, cells are seeded into 24 or 96-well plates, incubated at 37°C for 24 hours and treated with DEAE-Dextran at 37°C for 30 minutes. Virus is serially diluted in 1% DMEM, added to the cells incubating in DEAE-Dextran, and allowed to incubate for 3 hours at 37°C after which an additional cell media is added to each well. Following a final 48-hour incubation at 37°C, cells are either fixed, stained using X-Gal to visualize β-galactosidase expressing blue foci or frozen-thawed three times to measure luciferase activity.

Figure 8. Sequence alignment of subtype C ancestral and consensus *env* genes. Alignment of the subtype C ancestral (bottom line) and consensus (top line) *env* sequences showing a 95.5% sequence homology; amino acid sequence differences are indicated. One noted difference is the addition of a glycosylation site in the C ancestral *env* gene at the base of the V1 loop. A plus sign indicates a

within-class difference of amino acid at the indicated position; a bar indicates a change in the class of amino acid. Potential N-glycosylation sites are marked in blue. The position of truncation for the gp140 gene is also shown.

Figure 9. Expression of subtype C ancestral and consensus envelopes in 293T cells. Plasmids containing codon-optimized *gp160, gp140, or gp120* subtype C ancestral and consensus genes were transfected into 293T cells, and protein expression was examined by Western Blot analysis of cell lysates. 48-hours post-transfection, cell lysates were collected, total protein content determined by the BCA protein assay, and 2 μg of total protein was loaded per lane on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with HIV-1 plasma from a subtype C infected patient.

Figures 10A and 10B. Fig. 10A. *Trans* complementation of *env*-deficient HIV-1 with codon-optimized subtype C ancestral and consensus gp160 and gp140. Plasmids containing codon-optimized, subtype C ancestral or consensus *gp160* or *gp140* genes were co-transfected into 293T cells with an HIV-1/SG3Δ*env* provirus. 48 hours post-transfection cell supernatants containing pseudotyped virus were harvested, clarified by centrifugation, filtered through at 0.2μM filter, and pelleted through a 20% sucrose cushion. Quantification of p24 in each

virus pellet was determined using the Coulter HIV-1 p24 antigen assay; 25ng of p24 was loaded per lane on a 4-20% SDS-PAGE gel for particles containing a codon-optimized envelope. 250ng of p24 was loaded per lane for particles generated by co-transfection of a rev-dependent wild-type subtype C 96ZAM651*env* gene. Differences in the amount of p24 loaded per lane were necessary to ensure visualization of the rev-dependent envelopes by Western Blot. Proteins were transferred to a PVDF membrane and probed with pooled plasma from HIV-1 subtype B and subtype C infected individuals. Fig. 10B. Infectivity of virus particles containing subtype C ancestral and consensus envelope glycoproteins. Infectivity of pseudotyped virus containing ancestral or consensus *gp160* or *gp140* envelope was determined using the JC53-BL assay. Sucrose cushion purified virus particles were assayed by the Coulter p24 antigen assay, and 5-fold serial dilutions of each pellet were incubated with DEAE-Dextran treated JC53-BL cells. Following a 48-hour incubation period, cells were fixed and stained to visualize β-galactosidase expressing cells. Infectivity is represented as infectious units per ng of p24 to normalize for differences in the concentration of the input pseudovirions.

Figure 11. Co-receptor usage of subtype C ancestral and consensus envelopes. Pseudotyped particles containing ancestral or consensus envelope

were incubated with DEAE-Dextran treated JC53-BL cells in the presence of AMD3100 (a specific inhibitor of CXCR4), TAK779 (a specific inhibitor of CCR5), or AMD3000+TAK779 to determine co-receptor usage. NL4.3, an isolate known to utilize CXCR4, and YU-2, a known CCR5-using isolate, were included as controls.

Figures 12A-12C. Neutralization sensitivity of subtype C ancestral and consensus envelope glycoproteins. Equivalent amounts of pseudovirions containing the ancestral, consensus or 96ZAM651 *gp160* envelopes (1,500 infectious units) were pre-incubated with a panel of plasma samples from HIV-1 subtype C infected patients and then added to the JC53-BL cell monolayer in 96-well plates. Plates were cultured for two days and luciferase activity was measured as an indicator of viral infectivity. Virus infectivity is calculated by dividing the luciferase units (LU) produced at each concentration of antibody by the LU produced by the control infection. The mean 50% inhibitory concentration (IC_{50}) and the actual % neutralization at each antibody dilution are then calculated for each virus. The results of all luciferase experiments are confirmed by direct counting of blue foci in parallel infections.

Figures 13A-13F. Protein expression of consensus subtype C Gag (Fig. 13A) and Nef (Fig.

13B) following transfection into 293T cells. Consensus subtype C Gag and Nef amino acid sequences are set forth in Figs. 13C and 13D, respectively, and encoding sequences are set forth in Figs. 13E and 13F, respectively.

Figures 14A-14C. Figs. 14A and 14B show the Cons Env amino acid sequence and encoding sequence, respectively. Fig. 14C shows expression of Group M consensus CONs Env proteins using an *in vitro* transcription and translation system.

Figures 15A and 15B. Expression of CONs *env* gene in mammalian cells. (Fig. 15A - cell lysate, Fig. 15B - supernatant.)

Figures 16A and 16B. Infectivity (Fig. 16A) and coreceptor usage (Fig. 16B) of CON6 and CONs *env* genes.

Figures 17A-17C. Env protein incorporation in CON6 and CONs Env-pseudovirions. (Fig. 17A - lysate, Fig. 17B - supernatant, Fig. 17C pellet.)

Figures 18A-18D. Figs. 18A and 18B show subtype A consensus Env amino acid sequence and nucleic acid sequence encoding same, respectively. Figs. 18C and 18D show expression of A.con *env* gene in mammalian cells (Fig. 18C - cell lysate, Fig. 18D - supernatant).

Figures 19A-19H. M.con.gag (Fig. 19A), M.con.pol (Fig. 19B), M.con.nef (Fig. 19C) and C.con.pol (Fig. 19D) nucleic acid sequences and corresponding encoded amino acid sequences (Figs. 19E-19H, respectively).

Figures 20A-20D. Subtype B consensus *gag* (Fig. 20A) and *env* (Fig.20B) genes. Corresponding amino acid sequences are shown in Figs. 20C and 20D.

Figure 21. Expression of subtype B consensus *env* and *gag* genes in 293T cells. Plasmids containing codon-optimized subtype B consensus *gp160*, *gp140*, and *gag* genes were transfected into 293T cells, and protein expression was examined by Western Blot analysis of cell lysates. 48-hours post-transfection, cell lysates were collected, total protein content determined by the BCA protein assay, and 2 μg of total protein was loaded per lane on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with serum from an HIV-1 subtype B infected individual.

Figure 22. Co-receptor usage of subtype B consensus envelopes. Pseudotyped particles containing the subtype B consensus gp160 Env were incubated with DEAE-Dextran treated JC53-BL cells in the presence of AMD3100 (a specific inhibitor of CXCR4), TAK779 (a specific inhibitor of CCR5), and AMD3000+TAK779 to determine co-receptor usage.

NL4.3, an isolate known to utilize CXCR4 and YU-2, a known CCR5-using isolate, were included as controls.

Figures 23A and 23B. *Trans* complementation of *env*-deficient HIV-1 with codon-optimized subtype B consensus *gp160* and *gp140* genes. Plasmids containing codon-optimized, subtype B consensus *gp160* or *gp140* genes were co-transfected into 293T cells with an HIV-1/SG3Δenv provirus. 48-hours post-transfection cell supernatants containing pseudotyped virus were harvested, clarified in a tabletop centrifuge, filtered through a 0.2μM filter, and pellet through a 20% sucrose cushion. Quantification of p24 in each virus pellet was determined using the Coulter HIV-1 p24 antigen assay; 25 ng of p24 was loaded per lane on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with anti-HIV-1 antibodies from infected HIV-1 subtype B patient serum. *Trans* complementation with a rev-dependent NL4.3 *env* was included for control. Figure 23B. Infectivity of virus particles containing the subtype B concensus envelope. Infectivitiy of pseudotyped virus containing consensus B gp160 or gp140 was determined using the JC53-BL assay. Sucrose cushion purified virus particles were assayed by the Coulter p24 antigen assay, and 5-fold serial dilutions of each pellet were incubated with DEAE-Dextran treated JC53-BL cells. Following a 48-hour incubation period, cells were fixed and stained to visualize β-

galactosidase expressing cells. Infectivity is expressed as infectious units per ng of p24.

Figures 24A-24D. Neutralization sensitivity of virions containing subtype B consensus gp160 envelope. Equivalent amounts of pseudovirions containing the subtype B consensus or NL4.3 Env (gp160) (1,500 infectious units) were preincubated with three different monoclonal neutralizing antibodies and a panel of plasma samples from HIV-1 wubtype B infected individuals, and then added to the JC53-BL cell monolayer in 96-well plates. Plates were cultured for two days and luciferase activity was measured as an indicator of viral infectivity. Virus infectivity was calculated by dividing the luciferase units (LU) produced at each concentration of antibody by the LU produced by the control infection. The mean 50% inhibitory concentration (IC_{50}) and the actual % neutralization at each antibody dilution were then calculated for each virus. The results of all luciferase experiments were confirmed by direct counting of blue foci in parallel infections. Fig. 24A. Neutralization of Pseudovirions containing Subtype B consensus Env (gp160). Fig. 24B. Neutralization of Pseudovirions containing NL4.3 Env (gp160). Fig. 24C. Neutralization of Pseudovirions containing Subtype B consensus Env (gp160). Fig. 24D. Neutralization of Pseudovirions containing NL4.3 Env (gp160).

Figures 25A and 25B. Fig. 25A. Density and p24 analysis of sucrose gradient fractions. 0.5ml fractions were collected from a 20-60% sucrose gradient. Fraction number 1 represents the most dense fraction taken from the bottom of the gradient tube. Density was measured with a refractometer and the amount of p24 in each fraction was determined by the Coulter p24 antigen assay. Fractions 6-9, 10-15, 16-21, and 22-25 were pooled together and analyzed by Western Blot. As expected, virions sedimented at a density of 1.16-1.18 g/ml. Fig. 25B. VLP production by co-transfection of subtype B consensus *gag* and *env* genes. 293T cells were co-transfected with subtype B consensus *gag* and *env* genes. Cell supernatants were harvested 48-hours post-transfection, clarified through at 20% sucrose cushion, and further purified through a 20-60% sucrose gradient. Select fractions from the gradient were pooled, added to 20ml of PBS, and centrifuged overnight at 100,000 x g. Resuspended pellets were loaded onto a 4-20% SDS-PAGE gel, proteins were transferred to a PVDF membrane, and probed with plasma from an HIV-1 subtype B infected individual.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to an immunogen that induces antibodies that neutralize a wide spectrum of human immunodeficiency virus (HIV) primary isolates and/or that induces a T cell

response. The immunogen comprises at least one consensus or ancestral immunogen (e.g., Env, Gag, Nef or Pol), or portion or variant thereof. The invention also relates to nucleic acid sequences encoding the consensus or ancestral immunogen, or portion or variant thereof. The invention further relates to methods of using both the immunogen and the encoding sequences. While the invention is described in detail with reference to specific consensus and ancestral immunogens (for example, to a group M consensus Env), it will be appreciated that the approach described herein can be used to generate a variety of consensus or ancestral immunogens (for example, envelopes for other HIV-1 groups (e.g., N and O)).

In accordance with one embodiment of the invention, a consensus *env* gene can be constructed by generating consensus sequences of *env* genes for each subtype of a particular HIV-1 group (group M being classified into subtypes A-D, F-H, J an K), for example, from sequences in the Los Alamos HIV Sequence Database (using, for example, MASE (Multiple Aligned Sequence Editor)). A consensus sequence of all subtype consensuses can then be generated to avoid heavily sequenced subtypes (Gaschen et al, Science 296:2354-2360 (2002), Korber et al, Science 288:1789-1796 (2000)). In the case of the group M consensus *env* gene described in Example 1 (designated CON6), five highly variable regions from a CRF08_BC recombinant strain (98CN006) (V1, V2, V4, V5 and a region in cytoplasmic domain

of gp41) are used to fill in the missing regions in the sequence (see, however, corresponding regions for CONs). For high levels of expression, the codons of consensus or ancestral genes can be optimized based on codon usage for highly expressed human genes (Haas et al, Curr. Biol. 6:315-324 (2000), Andre et al, J. Virol. 72:1497-1503 (1998)).

It will be appreciated that the invention includes portions and variants of the sequences specifically disclosed herein. For example, and as regards the consensus and ancestral envelope sequences, the invention encompasses envelope sequences devoid of V3. Alternatively, V3 sequences can be selected from preferred sequences, for example, those described in U.S. Application No. 10/431,596 and U.S. Provisional Application No. 60/471,327.

A consensus or ancestral envelope of the invention can be been "activated" to expose intermediate conformations of neutralization epitopes that normally are only transiently or less well exposed on the surface of the HIV virion. The immunogen can be a "frozen" triggered form of a consensus or ancestral envelope that makes available specific epitopes for presentation to B lymphocytes. The result of this epitope presentation is the production of antibodies that broadly neutralize HIV. (Attention is directed to WO 02/024149 and to the activated/triggered envelopes described therein.)

The concept of a fusion intermediate immunogen is consistent with observations that the gp41 HR-2 region peptide, DP178, can capture an uncoiled conformation of gp41 (Furata et al, Nature Struct. Biol. 5:276 (1998)), and that formalin-fixed HIV-infected cells can generate broadly neutralizing antibodies (LaCasse et al, Science 283:357 (1997)). Recently a monoclonal antibody against the coiled-coil region bound to a conformational determinant of gp41 in HR1 and HR2 regions of the coiled-coil gp41 structure, but did not neutralize HIV (Jiang et al, J. Virol. 10213 (1998)). However, this latter study proved that the coiled-coil region is available for antibody to bind if the correct antibody is generated.

The immunogen of one aspect of the invention comprises a consensus or ancestral envelope either in soluble form or anchored, for example, in cell vesicles or in liposomes containing translipid bilayer envelope. To make a more native envelope, gp140 or gp160 consensus or ancestral sequences can be configured in lipid bilayers for native trimeric envelope formation. Alternatively, triggered gp160 in aldrithio 1-2 inactivated HIV-1 virions can be used as an immunogen. The gp160 can also exist as a recombinant protein either as gp160 or gp140 (gp140 is gp160 with the transmembrane region and possibly other gp41 regions deleted). Bound to gp160 or gp140 can be recombinant CCR5 or CXCR4 co-receptor proteins (or their extracellular domain peptide or protein fragments) or antibodies or other ligands

that bind to the CXCR4 or CCR5 binding site on gp120, and/or soluble CD4, or antibodies or other ligands that mimic the binding actions of CD4. Alternatively, vesicles or liposomes containing CD4, CCR5 (or CXCR4), or soluble CD4 and peptides reflective of CCR5 or CXCR4 gp120 binding sites. Alternatively, an optimal CCR5 peptide ligand can be a peptide from the N-terminus of CCR5 wherein specific tyrosines are sulfated (Bormier et al, Proc. Natl. Acad. Sci. USA 97:5762 (2001)). The triggered immunogen may not need to be bound to a membrane but may exist and be triggered in solution. Alternatively, soluble CD4 (sCD4) can be replaced by an envelope (gp140 or gp160) triggered by CD4 peptide mimetopes (Vitra et al, Proc. Natl. Acad. Sci. USA 96:1301 (1999)). Other HIV co-receptor molecules that "trigger" the gp160 or gp140 to undergo changes associated with a structure of gp160 that induces cell fusion can also be used. Ligation of soluble HIV gp140 primary isolate HIV 89.6 envelope with soluble CD4 (sCD4) induced conformational changes in gp41.

In one embodiment, the invention relates to an immunogen that has the characteristics of a receptor (CD4)-ligated consensus or ancestral envelope with CCR5 binding region exposed but unlike CD4-ligated proteins that have the CD4 binding site blocked, this immunogen has the CD4 binding site exposed (open). Moreover, this immunogen can be devoid of host CD4, which avoids the production of potentially

harmful anti-CD4 antibodies upon administration to a host.

The immunogen can comprise consensus or ancestral envelope ligated with a ligand that binds to a site on gp120 recognized by an A32 monoclonal antibodies (mab) (Wyatt et al, J. Virol. 69:5723 (1995), Boots et al, AIDS Res. Hum. Retro. 13:1549 (1997), Moore et al, J. Virol. 68:8350 (1994), Sullivan et al, J. Virol. 72:4694 (1998), Fouts et al, J. Virol. 71:2779 (1997), Ye et al, J. Virol. 74:11955 (2000)). One A32 mab has been shown to mimic CD4 and when bound to gp120, upregulates (exposes) the CCR5 binding site (Wyatt et al, J. Virol. 69:5723 (1995)). Ligation of gp120 with such a ligand also upregulates the CD4 binding site and does not block CD4 binding to gp120. Advantageously, such ligands also upregulate the HR-2 binding site of gp41 bound to cleaved gp120, uncleaved gp140 and cleaved gp41, thereby further exposing HR-2 binding sites on these proteins - each of which are potential targets for anti-HIV neutralizing antibodies.

In a specific aspect of this embodiment, the immunogen comprises soluble HIV consensus or ancestral gp120 envelope ligated with either an intact A32 mab, a Fab2 fragment of an A32 mab, or a Fab fragment of an A32 mab, with the result that the CD4 binding site, the CCR5 binding site and the HR-2 binding site on the consensus or ancestral envelope are exposed/upregulated. The immunogen can comprise consensus or ancestral envelope with an A32 mab (or

fragment thereof) bound or can comprise consensus or ancestral envelope with an A32 mab (or fragment thereof) bound and cross-linked with a cross-linker such as .3% formaldehyde or a heterobifunctional cross-linker such as DTSSP (Pierce Chemical Company). The immunogen can also comprise uncleaved consensus or ancestral gp140 or a mixture of uncleaved gp140, cleaved gp41 and cleaved gp120. An A32 mab (or fragment thereof) bound to consensus or ancestral gp140 and/or gp120 or to gp120 non-covalently bound to gp41, results in upregulation (exposure) of HR-2 binding sites in gp41, gp120 and uncleaved gp140. Binding of an A32 mab (or fragment thereof) to gp120 or gp140 also results in upregulation of the CD4 binding site and the CCR5 binding site. As with gp120 containing complexes, complexes comprising uncleaved gp140 and an A32 mab (or fragment thereof) can be used as an immunogen uncross-linked or cross-linked with cross-linker such as .3% formaldehyde or DTSSP. In one embodiment, the invention relates to an immunogen comprising soluble uncleaved consensus or ancestral gp140 bound and cross linked to a Fab fragment of an A32 mab, optionally bound and cross-linked to an HR-2 binding protein.

The consensus or ancestral envelope protein triggered with a ligand that binds to the A32 mab binding site on gp120 can be administered in combination with at least a second immunogen comprising a second envelope, triggered by a ligand that binds to a site distinct from the A32 mab

binding site, such as the CCR5 binding site recognized by mab 17b. The 17b mab (Kwong et al, Nature 393:648 (1998) available from the AIDS Reference Repository, NIAID, NIH) augments sCD4 binding to gp120. This second immunogen (which can also be used alone or in combination with triggered immunogens other than that described above) can, for example, comprise soluble HIV consensus or ancestral envelope ligated with either the whole 17b mab, a Fab2 fragment of the 17b mab, or a Fab fragment of the 17b mab. It will be appreciated that other CCR5 ligands, including other antibodies (or fragments thereof), that result in the CD4 binding site being exposed can be used in lieu of the 17b mab. This further immunogen can comprise gp120 with the 17b mab, or fragment thereof, (or other CCR5 ligand as indicated above) bound or can comprise gp120 with the 17b mab, or fragment thereof, (or other CCR5 ligand as indicated above) bound and cross-linked with an agent such as .3% formaldehyde or a heterobifunctional cross-linker, such as DTSSP (Pierce Chemical Company). Alternatively, this further immunogen can comprise uncleaved gp140 present alone or in a mixture of cleaved gp41 and cleaved gp120. Mab 17b, or fragment thereof (or other CCR5 ligand as indicated above) bound to gp140 and/or gp120 in such a mixture results in exposure of the CD4 binding region. The 17b mab, or fragment thereof, (or other CCR5 ligand as indicated above) gp140 complexes can be present uncross-linked or

cross-linked with an agent such as .3% formaldehyde or DTSSP.

Soluble HR-2 peptides, such as T649Q26L and DP178, can be added to the above-described complexes to stabilize epitopes on consensus gp120 and gp41 as well as uncleaved consensus gp140 molecules, and can be administered either cross-linked or uncross-linked with the complex.

A series of monoclonal antibodies (mabs) have been made that neutralize many HIV primary isolates, including, in addition to the 17b mab described above, mab IgG1b12 that binds to the CD4 binding site on gp120(Roben et al, J. Virol. 68:482 (1994), Mo et al, J. Virol. 71:6869 (1997)), mab 2G12 that binds to a conformational determinant on gp120 (Trkola et al, J. Virol. 70:1100 (1996)), and mab 2F5 that binds to a membrane proximal region of gp41 (Muster et al, J. Virol. 68:4031 (1994)).

As indicated above, various approaches can be used to "freeze" fusogenic epitopes in accordance with the invention. For example, "freezing" can be effected by addition of the DP-178 or T-649Q26L peptides that represent portions of the coiled coil region, and that when added to CD4-triggered consensus or ancestral envelope, result in prevention of fusion (Rimsky et al, J. Virol. 72:986-993 (1998)). HR-2 peptide bound consensus or ancestral gp120, gp140, gp41 or gp160 can be used as an immunogen or crosslinked by a reagent such as DTSSP or DSP (Pierce Co.), formaldehyde or other crosslinking agent that has a similar effect.

"Freezing" can also be effected by the addition of 0.1% to 3% formaldehyde or paraformaldehyde, both protein cross-linking agents, to the complex, to stabilize the CD4, CCR5 or CXCR4, HR-2 peptide gp160 complex, or to stabilize the "triggered" gp41 molecule, or both (LaCasse et al, Science 283:357-362 (1999)).

Further, "freezing" of consensus or ancestral gp41 or gp120 fusion intermediates can be effected by addition of heterobifunctional agents such as DSP (dithiobis[succimidylproprionate]) (Pierce Co. Rockford, ILL., No. 22585ZZ) or the water soluble DTSSP (Pierce Co.) that use two NHS esters that are reactive with amino groups to cross link and stabilize the CD4, CCR5 or CXCR4, HR-2 peptide gp160 complex, or to stabilize the "triggered" gp41 molecule, or both.

Analysis of T cell immune responses in immunized or vaccinated animals and humans shows that the envelope protein is normally not a main target for T cell immune response although it is the only gene that induces neutralizing antibodies. HIV-1 Gag, Pol and Nef proteins induce a potent T cell immune response. Accordingly, the invention includes a repertoire of consensus or ancestral immunogens that can induce both humoral and cellular immune responses. Subunits of consensus or ancestral sequences can be used as T or B cell immunogens.

The immunogen of the invention can be formulated with a pharmaceutically acceptable

carrier and/or adjuvant (such as alum) using techniques well known in the art. Suitable routes of administration of the present immunogen include systemic (e.g. intramuscular or subcutaneous). Alternative routes can be used when an immune response is sought in a mucosal immune system (e.g., intranasal).

The immunogens of the invention can be chemically synthesized and purified using methods which are well known to the ordinarily skilled artisan. The immunogens can also be synthesized by well-known recombinant DNA techniques. Nucleic acids encoding the immunogens of the invention can be used as components of, for example, a DNA vaccine wherein the encoding sequence is administered as naked DNA or, for example, a minigene encoding the immunogen can be present in a viral vector. The encoding sequence can be present, for example, in a replicating or non-replicating adenoviral vector, an adeno-associated virus vector, an attenuated mycobacterium tuberculosis vector, a Bacillus Calmette Guerin (BCG) vector, a vaccinia or Modified Vaccinia Ankara (MVA) vector, another pox virus vector, recombinant polio and other enteric virus vector, Salmonella species bacterial vector, Shigella species bacterial vector, Venezuelean Equine Encephalitis Virus (VEE) vector, a Semliki Forest Virus vector, or a Tobacco Mosaic Virus vector. The encoding sequence, can also be expressed as a DNA plasmid with, for example, an active promoter such as a CMV promoter. Other live

vectors can also be used to express the sequences of the invention. Expression of the immunogen of the invention can be induced in a patient's own cells, by introduction into those cells of nucleic acids that encode the immunogen, preferably using codons and promoters that optimize expression in human cells. Examples of methods of making and using DNA vaccines are disclosed in U.S. Pat. Nos. 5,580,859, 5,589,466, and 5,703,055.

The composition of the invention comprises an immunologically effective amount of the immunogen of this invention, or nucleic acid sequence encoding same, in a pharmaceutically acceptable delivery system. The compositions can be used for prevention and/or treatment of immunodeficiency virus infection. The compositions of the invention can be formulated using adjuvants, emulsifiers, pharmaceutically-acceptable carriers or other ingredients routinely provided in vaccine compositions. Optimum formulations can be readily designed by one of ordinary skill in the art and can include formulations for immediate release and/or for sustained release, and for induction of systemic immunity and/or induction of localized mucosal immunity (e.g, the formulation can be designed for intranasal administration). The present compositions can be administered by any convenient route including subcutaneous, intranasal, oral, intramuscular, or other parenteral or enteral route. The immunogens can be administered as a single dose or multiple doses. Optimum immunization schedules

can be readily determined by the ordinarily skilled artisan and can vary with the patient, the composition and the effect sought.

The invention contemplates the direct use of both the immunogen of the invention and/or nucleic acids encoding same and/or the immunogen expressed as minigenes in the vectors indicated above. For example, a minigene encoding the immunogen can be used as a prime and/or boost.

Certain aspects of the invention can be described in greater detail in the non-limiting Examples that follows.

EXAMPLE 1

Artificial HIV-1 Group M Consensus Envelope

EXPERIMENTAL DETAILS

Expression of CON6 gp120 and gp140 proteins in recombinant vaccinia viruses (VV). To express and purify the secreted form of HIV-1 CON6 envelope proteins, CON6 gp120 and gp140CF plasmids were constructed by introducing stop codons after the gp120 cleavage site (REKR) and before the transmembrane domain (YIKIFIMIVGGLIGLRIVFAVLSIVN), respectively. The gp120/gp41 cleavage site and fusion domain of gp41 were deleted in the gp140CF protein. Both CON6 gp120 and gp140CF DNA constructs were cloned into the pSC65 vector (from Bernard Moss, NIH, Bethesda, MD) at SalI and KpnI

restriction enzyme sites. This vector contains the lacZ gene that is controlled by the p7.5 promoter. A back-to-back P E/L promoter was used to express CON6 *env* genes. BSC-1 cells were seeded at 2×10^5 in each well in a 6-well plate, infected with wild-type vaccinia virus (WR) at a MOI of 0.1 pfu/cell, and 2 hr after infection, pSC65-derived plasmids containing CON6 *env* genes were transfected into the VV-infected cells and recombinant (r) VV selected as described (Moss and Earl, Current Protocols in Molecular Biology, eds, Ausubel et al (John Wiley & Sons, Inc. Indianapolis, IN) pp. 16.15.1-16.19.9 (1998)). Recombinant VV that contained the CON6 *env* genes were confirmed by PCR and sequencing analysis. Expression of the CON6 envelope proteins was confirmed by SDS-PAGE and Western blot assay. Recombinant CON6 gp120 and gp140CF were purified with agarose *galanthus Nivalis* lectin beads (Vector Labs, Burlingame, CA), and stored at -70°C until use. Recombinant VV expressing JRFL (vCB-28) or 96ZM651 (vT241R) gp160 were obtained from the NIH AIDS Research and Reference Reagent Program (Bethesda, MD).

Monoclonal Antibodies and gp120 Wild-type Envelopes. Human mabs against a conformational determinant on gp120 (A32), the gp120 V3 loop (F39F) and the CCR5 binding site (17b) were the gifts of James Robinson (Tulane Medical School, New Orleans, LA) (Wyatt et al, Nature 393;705-711 (1998), Wyatt et al, J. Virol. 69:5723-5733 (1995)). Mabs 2F5,

447, b12, 2G12 and soluable CD4 were obtained from the NIH AIDS Research and Reference Reagent Program (Bethesda, MD) (Gorny e tal, J. Immunol. 159:5114-5122 (1997), Nyambi et al, J. Virol. 70:6235-6243 (1996), Purtscher et al, AIDS Res. Hum. Retroviruses 10:1651-1658 (1994), Trkola et al, J. Virol 70:1100-1108 (1996)). T8 is a murine mab that maps to the gp120 C1 region (a gift from P. Earl, NIH, Bethesda, MD). BaL (subtype B), 96ZM651 (subtype C), and 93TH975 (subtype E) gp120s were provided by QBI, Inc. and the Division of AIDS, NIH. CHO cell lines that express 92U037 (subtype A) and 93BR029 (subtype F) gp140 (secreted and uncleaved) were obtained from NICBS, England.

Surface Plasmon Resonance Biosensor (SPR) Measurements and ELISA. SPR biosensor measurements were determined on a BIAcore 3000 instrument (BIAcore Inc., Uppsala, Sweden) instrument and data analysis was performed using BIAevaluation 3.0 software (BIAcore Inc, Upsaala, Sweden). Anti-gp120 mabs (T8, A32, 17b, 2G12) or sCD4 in 10mM Na-acetate buffer, pH 4.5 were directly immobilized to a CM5 sensor chip using a standard amine coupling protocol for protein immobilization. FPLC purified CON6 gp120 monomer or gp140CF oligomer recombinant proteins were flowed over CM5 sensor chips at concentrations of 100 and 300 μg/ml, respectively. A blank in-line reference surface (activated and de-activated for amine coupling) or non-bonding mab controls were used to subtract non-specific or bulk

responses. Soluble 89.6 gp120 and irrelevant IgG was used as a positive and negative control respectively and to ensure activity of each mab surface prior to injecting the CON6 Env proteins. Binding of CON6 envelope proteins was monitored in real-time at 25°C with a continuous flow of PBS (150 mM NaCl, 0.005% surfactant P20), pH 7.4 at 10-30 µl/min. Bound proteins were removed and the sensor surfaces were regenerated following each cycle of binding by single or duplicate 5-10 µl pulses of regeneration solution (10 mM glycine-HCl, pH 2.9). ELISA was performed to determine the reactivity of various mabs to CON6 gp120 and gp140CF proteins as described (Haynes et al, AIDS Res. Hum. Retroviruses 11:211-221 (1995)). For assay of human mab binding to rgp120 or gp140 proteins, end-point titers were defined as the highest titer of mab (beginning at 20 µg/ml) at which the mab bound CON6 gp120 and gp140CF Env proteins 3 fold over background control (non-binding human mab).

Infectivity and coreceptor usage assays. HIV-1/SG3Δenv and CON6 or control *env* plasmids were cotransfected into human 293T cells. Pseudotyped viruses were harvested, filtered and p24 concentration was quantitated (DuPont/NEN Life Sciences, Boston, MA). Equal amounts of p24 (5 ng) for each pseudovirion were used to infect JC53-BL cells to determine the infectivity (Derdeyn e al, J. Virol. 74:8358-8367 (2000), Wei et al, Antimicrob

Agents Chemother. 46:1896-1905 (2002)). JC53-BL cells express CD4, CCR5 and CXCR4 receptors and contain a β-galactosidase (β-gal) gene stably integrated under the transcriptional control of an HIV-1 long terminal repeat (LTR). These cells can be used to quantify the infectious titers of pseudovirion stocks by staining for β-gal expression and counting the number of blue cells (infectious units) per microgram of p24 of pseudovirons (IU/μg p24) (Derdeyn e al, J. Virol. 74:8358-8367 (2000), Wei et al, Antimicrob Agents Chemother. 46:1896-1905 (2002)). To determine the coreceptor usage of the CON6 *env* gene, JC53BL cells were treated with 1.2 μM AMD3100 and 4 μM TAK-799 for 1 hr at 37°C then infected with equal amounts of p24 (5 ng) of each Env pseudotyped virus. The blockage efficiency was expressed as the percentage of the infectious units from blockage experiments compared to that from control culture without blocking agents. The infectivity from control group (no blocking agent) was arbitrarily set as 100%.

Immunizations. All animals were housed in the Duke University Animal Facility under AALAC guidelines with animal use protocols approved by the Duke University Animal Use and Care Committee. Recombinant CON6 gp120 and gp140CF glycoproteins were formulated in a stable emulsion with RIBI-CWS adjuvant based on the protocol provided by the manufacturer (Sigma Chemical Co., St. Louis, MO).

For induction of anti-envelope antibodies, each of four out-bred guinea pigs (Harlan Sprague, Inc., Chicago, IL) was given 100 µg either purified CON6 gp120 or gp140CF subcutaneously every 3 weeks (total of 5 immunizations). Serum samples were heat-inactivated (56°C, 1 hr), and stored at -20°C until use.

For induction of anti-envelope T cell responses, 6-8 wk old female BALB/c mice (Frederick Cancer Research and Developmental Center, NCI, Frederick, MD) were immunized i.m. in the quadriceps with 50 µg plasmid DNA three times at a 3-week interval. Three weeks after the last DNA immunization, mice were boosted with 10^7 PFU of rVV expressing Env proteins. Two weeks after the boost, all mice were euthanized and spleens were removed for isolation of splenocytes.

Neutralization assays. Neutralization assays were performed using either a MT-2 assay as described in Bures et al, AIDS Res. Hum. Retroviruses 16:2019-2035 (2000), a luciferase-based multiple replication cycle HIV-1 infectivity assay in 5.25.GFP.Luc.M7 cells using a panel of HIV-1 primary isolates (Bures et al, AIDS Res. Hum. Retroviruses 16:2019-2035 (2000), Bures et al, J. Virol. 76:2233-2244 (2002)), or a syncytium (fusion from without) inhibition assay using inactivated HIV-1 virions (Rossio et al, J. Virol. 72:7992-8001 (1998)). In the luciferase-based assay,

neutralizing antibodies were measured as a function of a reduction in luciferase acitivity in 5.25.EGFP.Luc.M7 cells provided by Nathaniel R. Landau, Salk Institute, La Jolla, CA (Brandt et al, J. Biol. Chem. 277:17291-17299 (2002)). Five hundred tissue culture infectious dose 50 ($TCID_{50}$) of cell-free virus was incubated with indicated serum dilutions in 150 µl (1 hr, at 37°C) in triplicate in 96-well flat-bottom culture plates. The 5.25.EGFP.Luc.M7 cells were suspended at a density of 5 x 10^5/ml in media containing DEAE dextran (10 µg/ml). Cells (100 µl) were added and until 10% of cells in control wells (no test serum sample) were positive for GFP expression by fluorescence microscopy. At this time the cells were concentrated 2-fold by removing one-half volume of media. A 50 µl suspension of cells was transferred to 96-well white solid plates (Costar, Cambridge, MA) for measurement of luciferase activity using Bright-Glo™ substrate (Promega, Madison, WI) on a Wallac 1420 Multilabel Counter (PerkinElmer Life Sciences, Boston, MA). Neutralization titers in the MT-2 and luciferase assays were those where $\geq$ 50% virus infection was inhibited. Only values that titered beyond 1:20 (i.e. >1:30) were considered significantly positive. The syncytium inhibition "fusion from without" assay utilized HIV-1 aldrithiol-2 (AT-2) inactivated virions from HIV-1 subtype B strains ADA and AD8 (the gift of Larry Arthur and Jeffrey Lifson, Frederick Research Cancer

Facility, Frederick, MD) added to SupT1 cells, with syncytium inhibition titers determined as those titers where ≥90% of syncytia were inhibited compared to prebleed sera.

Enzyme linked immune spot (ELISPOT) assay. Single-cell suspensions of splenocytes from individual immunized mice were prepared by mincing and forcing through a 70 μm Nylon cell strainer (BD Labware, Franklin Lakes, NJ). Overlapping Env peptides of CON6 gp140 (159 peptides, 15mers overlapping by 11) were purchased from Boston Bioscence, Inc (Royal Oak, MI). Overlapping Env peptides of MN gp140 (subtype B; 170 peptides, 15mers overlapping by 11) and Chn19 gp140 (subtype C; 69 peptides, 20mers overlapping by 10) were obtained from the NIH AIDS Research and Reference Reagent Program (Bethesda, MD). Splenocytes (5 mice/group) from each mouse were stimulated *in vitro* with overlapping Env peptides pools from CON6, subtype B and subtype C Env proteins. 96-well PVDF plates (MultiScreen-IP, Millipore, Billerica, MA) were coated with anti-IFN-γ mab (5 μg/ml, AN18; Mabtech, Stockholm, Sweden). After the plates were blocked at 37 C for 2 hr using complete Hepes buffered RPMI medium, 50μl of the pooled overlapping envelope peptides (13 CON6 and MN pools, 13-14 peptides in each pool; 9 Chn19 pool, 7-8 peptide in each pool) at a final concentration of 5 μg/ml of each were added to the plate. Then 50 μl of

splenocytes at a concentration of 1.0 X 10^7/ml were added to the wells in duplicate and incubated for 16 hr at 37 C with 5% CO_2. The plates were incubated with 100 µl of a 1:1000 dilution of streptavidin alkaline phosphatase (Mabtech, Stockholm, Sweden), and purple spots developed using 100 µl of BCIP/NBT (Plus) Alkaline Phosphatase Substrate (Moss, Pasadena, MD). Spot forming cells (SFC) were measured using an Immunospot counting system (CTL Analyzers, Cleveland, OH). Total responses for each envelope peptide pool are expressed as SFCs per 10^6 splenocytes.

RESULTS

CON6 Envelope Gene Design, Construction and Expression. An artificial group M consensus *env* gene (CON6) was constructed by generating consensus sequences of *env* genes for each HIV-1 subtype from sequences in the Los Alamos HIV Sequence Database, and then generating a consensus sequence of all subtype consensuses to avoid heavily sequenced subtypes (Gaschen et al, Science 296:2354-2360 (2002), Korber et al, Science 288:1789-1796 (2000)). Five highly variable regions from a CRF08_BC recombinant strain (98CN006) (V1, V2, V4, V5 and a region in cytoplasmic domain of gp41) were then used to fill in the missing regions in CON6 sequence. The CON6 V3 region is group M consensus (Figure 1A). For high levels of expression, the codons of CON6 *env* gene were optimized based on codon usage for

highly expressed human genes (Haas et al, Curr. Biol. 6:315-324 (2000), Andre et al, J. Virol. 72:1497-1503 (1998)). (See Fig. 1D.) The codon optimized CON6 env gene was constructed and subcloned into pcDNA3.1 DNA at EcoR I and BamH I sites (Gao et al, AIDS Res. Hum. Retroviruses, 19:817-823 (2003)). High levels of protein expression were confirmed with Western-blot assays after transfection into 293T cells. To obtain recombinant CON6 Env proteins for characterization and use as immunogens, rVV was generated to express secreted gp120 and uncleaved gp140CF (Figure 1B). Purity for each protein was $\geq$90% as determined by Coomassie blue gels under reducing conditions (Figure 1C).

CD4 Binding Domain and Other Wild-type HIV-1 Epitopes are Preserved on CON6 Proteins. To determine if CON6 proteins can bind to CD4 and express other wild-type HIV-1 epitopes, the ability of CON6 gp120 and gp140CF to bind soluble(s) CD4, to bind several well-characterized anti-gp120 mabs, and to undergo CD4-induced conformational changes was assayed. First, BIAcore CM5 sensor chips were coated with either sCD4 or mabs to monitor their binding activity to CON6 Env proteins. It was found that both monomeric CON6 gp120 and oligomeric gp140CF efficiently bound sCD4 and anti-gp120 mabs T8, 2G12 and A32, but did not constitutively bind mab 17b, that recognizes a CD4 inducible epitope in the CCR5 binding site of gp120 (Figures 2A and 2B).

Both sCD4 and A32 can expose the 17b binding epitope after binding to wild-type gp120 (Wyatt et al, Nature 393;705-711 (1998), Wyatt et al, J. Virol. 69:5723-5733 (1995)). To determine if the 17b epitope could be induced on CON6 Envs by either sCD4 or A32, sCD4, A32 and T8 were coated on sensor chips, then CON6 gp120 or gp140CF captured, and mab 17b binding activity monitored. After binding sCD4 or mab A32, both CON6 gp120 and gp140CF were triggered to undergo conformational changes and bound mab 17b (Figures 2C and 2D). In contrast, after binding mab T8, the 17b epitope was not exposed (Figures 2C and 2D). ELISA was next used to determine the reactivity of a panel of human mabs against the gp120 V3 loop (447, F39F), the CD4 binding site (b12), and the gp41 neutralizing determinant (2F5) to CON6 gp120 and gp140CF (Figure 2E). Both CON6 rgp120 and rgp140CF proteins bound well to neutralizing V3 mabs 447 and F39F and to the potent neutralizing CD4 binding site mab b12. Mab 2F5, that neutralizes HIV-1 primary isolates by binding to a C-terminal gp41 epitope, also bound well to CON6 gp140CF (Figure 2E).

CON6 env Gene is Biologically Functional and Uses CCR5 as its Coreceptor. To determine whether CON6 envelope gene is biologically functional, it was co-transfected with the *env*-defective SG3 proviral clone into 293T cells. The pseudotyped viruses were harvested and JC53BL cells infected. Blue cells were detected in JC53-BL cells infected

with the CON6 Env pseudovirions, suggesting that CON6 Env protein is biologically functional (Figure 3A). However, the infectious titers were 1-2 logs lower than that of pseudovirions with either YU2 or NL4-3 wild-type HIV-1 envelopes.

The co-receptor usage for the CON6 *env* gene was next determined. When treated with CXCR4 blocking agent AMD3100, the infectivity of NL4-3 Env-pseudovirons was blocked while the infectivity of YU2 or CON6 Env-pseudovirons was not inhibited (Figure 3B). In contrast, when treated with CCR5 blocking agent TAK-779, the infectivity of NL4-3 Env-pseudovirons was not affected, while the infectivity of YU2 or CON6 Env-pseudovirons was inhibited. When treated with both blocking agents, the infectivity of all pseudovirions was inhibited. Taken together, these data show that the CON6 envelope uses the CCR5 co-receptor for its entry into target cells.

Reaction of CON6 gp120 With Different Subtype Sera. To determine if multiple subtype linear epitopes are preserved on CON6 gp120, a recombinant Env protein panel (gp120 and gp140) was generated. Equal amounts of each Env protein (100 ng) were loaded on SDS-polyacrylamide gels, transferred to nitrocellulose, and reacted with subtype A through G patient sera as well as anti-CON6 gp120 guinea pig sera (1:1,000 dilution) in Western blot assays. For each HIV-1 subtype, four to six patient sera were

tested. One serum representative for each subtype is shown in Figure 4.

It was found that whereas all subtype sera tested showed variable reactivities among Envs in the panel, all group M subtype patient sera reacted equally well with CON6 gp120 Env protein, demonstrating that wild-type HIV-1 Env epitopes recognized by patient sera were well preserved on the CON6 Env protein. A test was next made as to whether CON6 gp120 antiserum raised in guinea pigs could react to different subtype Env proteins. It was found that the CON6 serum reacted to its own and other subtype Env proteins equally well, with the exception of subtype A Env protein (Figure 4).

Induction of T Cell Responses to CON6, Subtype B and Subtype C Envelope Overlapping Peptides. To compare T cell immune responses induced by CON6 Env immunogens with those induced by subtype specific immunogens, two additional groups of mice were immunized with subtype B or subtype C DNAs and with corresponding rVV expressing subtype B or C envelope proteins. Mice immunized with subtype B (JRFL) or subtype C (96ZM651) Env immunogen had primarily subtype-specific T cell immune responses (Figure 5). IFN-γ SFCs from mice immunized with JRFL (subtype B) immunogen were detected after stimulation with subtype B (MN) peptide pools, but not with either subtype C (Chn19) or CON6 peptide pools. IFN-γ SFCs from mice immunized with 96ZM651 (subtype C)

immunogen were detected after the stimulation with both subtype C (Chn19) and CON6 peptide pools, but not with subtype B (MN) peptide pools. In contrast, IFN-γ SFCs were identified from mice immunized with CON6 Env immunogens when stimulated with either CON6 peptide pools as well as by subtype B or C peptide pools (Figure 5). The T cell immune responses induced by CON6 gp140 appeared more robust than those induced by CON6 gp120. Taken together, these data demonstrated that CON6 gp120 and gp140CF immunogens were capable of inducing T cell responses that recognized T cell epitopes of wild-type subtype B and C envelopes.

Induction of Antibodies by Recombinant CON6 gp120 and gp140CF Envelopes that Neutralize HIV-1 Subtype B and C Primary Isolates. To determine if the CON6 envelope immunogens can induce antibodies that neutralize HIV-1 primary isolates, guinea pigs were immunized with either CON6 gp120 or gp140CF protein. Sera collected after 4 or 5 immunizations were used for neutralization assays and compared to the corresponding prebleed sera. Two AT-2 inactivated HIV-1 isolates (ADA and AD8) were tested in syncytium inhibition assays (Table 1A). Two subtype B SHIV isolates, eight subtype B primary isolates, four subtype C, and one each subtype A, D, and E primary isolates were tested in either the MT-2 or the luciferase-based assay (Table 1B). In the syncytium inhibition assay, it was found that

antibodies induced by both CON 6 gp120 and gp140CF proteins strongly inhibited AT-2 inactivated ADA and AD8-induced syncytia (Table 1A). In the MT-2 assay, weak neutralization of 1 of 2 SHIV isolates (SHIV SF162P3) by two gp120 and one gp140CF sera was found (Table 1B). In the luciferase-based assay, strong neutralization of 4 of 8 subtype B primary isolates (BXO8, SF162, SS1196, and BAL) by all gp120 and gp140CF sera was found, and weak neutralization of 2 of 8 subtype B isolates (6101, 0692) by most gp120 and gp140CF sera was found. No neutralization was detected against HIV-1 PAVO (Table 1B). Next, the CON6 anti-gp120 and gp140CF sera were tested against four subtype C HIV-1 isolates, and weak neutralization of 3 of 4 isolates (DU179, DU368, and S080) was found, primarily by anti-CON6 gp120 sera. One gp140CF serum, no. 653, strongly neutralized DU179 and weakly neutralized S080 (Table 1B). Finally, anti-CON6 Env sera strongly neutralized a subtype D isolate (93ZR001), weakly neutralized a subtype E (CM244) isolate, and did not neutralize a subtype A (92RW020) isolate.

Table 1A

Ability of HIV-1 Group M Consensus Envel pe CON6 Pr t ins to Induce Fusion Inhibiting Antibodies

Guinea Pig No.	Immunogen	Syncytium Inhibition antibody titer[1] AD8	ADA
646	gp120	270	270
647	gp120	90	90
648	gp120	90	270
649	gp120	90	90
Geometric Mean Titer		119	156
650	gp140	270	270
651	gp140	90	90
652	gp140	810	810
653	gp140	270	90
Geometric Mean Titer		270	207

[1]Reciprocal serum dilution at which HIV-induced syncytia of Sup T1 cells was inhibited by >90% compared to pre-immune serum. All prebleed sera were negative (titer <10).

Table 1B

Ability of Group M Consensus HIV-1 Envelope CON6 gp120 and gp140CF Proteins to Induce Antibodies that Neutralize HIV Primary Isolates

HIV Isolate (Subtype)	CON6 gp120 Protein Guinea Pig No.					CON6 gp140CF Protein Guinea Pig No.					Controls		
	646	647	648	649	GMT	650	651	652	653	GMT	$TriMab_2$	CD4-IgG2	HIV+ Serum
SHIV 89.6P*(B)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	NT	NT	NT
SHIV SF162P3*(B)	<20	30	48	<20	<20	27	<20	<20	<20	<20	NT	0.2µg/ml	NT
BX08(B)	270	183	254	55	102	199	64	229	150	187	0.7µg/ml	NT	2384
6101(B)	<20	38	35	<20	<20	<20	90	72	73	39	1.1µg/ml	NT	NT
BG1168(B)	<20	<20	<20	<20	<20	40	<20	<20	25	<20	2.7µg/ml	NT	NT
0692(B)	31	32	34	<20	24	28	33	30	45	33	0.8µg/ml	NT	769
PAVO(B)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	2.9µg/ml	NT	NT
SF162(B)	2,146	308	110	282	379	206	5,502	15,098	174	1,313	NT	NT	>540
SS1196(B)	206	26	148	59	83	381	401	333	81	253	NT	NT	301#
BAL(B)	123	90	107	138	113	107	146	136	85	116	NT	NT	3307
92RW020(A)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	NT	NT	693
DU179(C)	<20	43	<20	24	<20	<20	<20	24	515	33	NT	0.8µg/ml	NT
DU368(C)	25	35	62	<20	27	<20	<20	<20	23	<20	NT	2.3µg/ml	NT
S021(C)	<20	<20	33	<20	<20	<20	<20	<20	<20	<20	NT	8.3µg/ml	NT
S080(C)	24	37	70	41	40	<20	<20	<20	52	<20	NT	3.4µg/ml	NT
93ZR001(D)	275	144	126	114	154	306	195	129	173	191	NT	NT	693
CM244(E)	35	43	64	ND	46	31	25	27	25	26	NT	NT	693

*MT-2 Assay; All other HIV isolates were tested in the M7-luciferase assay.

HIV-1 isolates QH0692, SS1196, SF162, 6101, BX08, BG1168, BAL were assayed with post-injection 5 serum; other HIV-1 isolates were assayed with post-injection 4 serum. ND = not done.

HIV+ sera was either HIV-1+ human serum (LEH3) or an anti-gp120 guinea pig serum (#) with known neutralizing activity for HIV-1 isolate SS1196. GMT = geometric mean titer of four animals per group. Neutralizing titers reported are after subtraction of any background neutralization in prebleed sera.

$TriMab_2$ = a mixture of human mabs 2F5, b12, 2G12.

CONCLUSIONS

The production of an artificial HIV-1 Group M consensus *env* gene (CON6) has been described that encodes a functional Env protein that is capable of utilizing the CCR5 co-receptor for mediating viral entry. Importantly, this Group M consensus envelope gene could induce T and B cell responses that recognized epitopes of subtype B and C HIV-1 primary isolates.

The correlates of protection to HIV-1 are not conclusively known. Considerable data from animal models and studies in HIV-1-infected patients suggest the goal of HIV-1 vaccine development should be the induction of broadly-reactive CD4+ and CD8+ anti-HIV-1 T cell responses (Letvin et al, Annu. Rev. Immunol. 20:73-99 (2002)) and high levels of antibodies that neutralize HIV-1 primary isolates of multiple subtypes (Mascola et al, J. Virol. 73:4009-4018 (1999), Mascola et al, Nat. Med. 6:270-210 (2000)).

The high level of genetic variability of HIV-1 has made it difficult to design immunogens capable of inducing immune responses of sufficient breadth to be clinically useful. Epitope based vaccines for T and B cell responses (McMichael et al, Vaccine 20:1918-1921 (2002), Sbai et al, Curr. Drug Targets Infect, Disord. 1:303-313 (2001), Haynes, Lancet 348:933-937 (1996)), constrained envelopes reflective of fusion intermediates (Fouts et al, Proc. Natl. Acad. Sci. USA 99:11842-22847 (2002)),

as well as exposure of conserved high-order structures for induction of anti-HIV-1 neutralizing antibodies have been proposed to overcome HIV-1 variability (Roben et al, J. Virol. 68:4821-4828 (1994), Saphire et al, Science 293:1155-1159 (2001)). However, with the ever-increasing diversity and rapid evolution of HIV-1, the virus is a rapidly moving complex target, and the extent of complexity of HIV-1 variation makes all of these approaches problematic. The current most common approach to HIV-1 immunogen design is to choose a wild-type field HIV-1 isolate that may or may not be from the region in which the vaccine is to be tested. Polyvalent envelope immunogens have been designed incorporating multiple envelope immunogens (Bartlett et al, AIDS 12:1291-1300 (1998), Cho et al, J. Virol. 75:2224-2234 (2001)).

The above-described study tests a new strategy for HIV-1 immunogen design by generating a group M consensus *env* gene (CON6) with decreased genetic distance between this candidate immunogen and wild-type field virus strains. The CON6 *env* gene was generated for all subtypes by choosing the most common amino acids at most positions (Gaschen et al, Science 296:2354-2360 (2002), Korber et al, Science 288:1789-1796 (2000)). Since only the most common amino acids were used, the majority of antibody and T cell epitopes were well preserved. Importantly, the genetic distances between the group M consensus *env* sequence and any subtype *env* sequences was about 15%, which is only half of that between wild-type

subtypes (30%) (Gaschen et al, Science 296:2354-2360 (2002)). This distance is approximately the same as that among viruses within the same subtype. Further, the group M consensus *env* gene was also about 15% divergent from any recombinant viral *env* gene, as well, since CRFs do not increase the overall genetic divergence among subtypes.

Infectivity of CON6-Env pseudovirions was confirmed using a single-round infection system, although the infectivity was compromised, indicating the artificial envelope was not in an "optimal" functional conformation, but yet was able to mediate virus entry. That the CON6 envelope used CCR5 (R5) as its coreceptor is important, since majority of HIV-1 infected patients are initially infected with R5 viruses.

BIAcore analysis showed that both CON6 gp120 and gp140CF bound sCD4 and a number of mabs that bind to wild-type HIV-1 Env proteins. The expression of the CON6 gp120 and 140CF proteins that are similar antigenically to wild-type HIV-1 envelopes is an important step in HIV-1 immunogen development. However, many wild-type envelope proteins express the epitopes to which potent neutralizing human mabs bind, yet when used as immunogens themselves, do not induce broadly neutralizing anti-HIV-1 antibodies of the specificity of the neutralizing human mabs.

The neutralizing antibody studies were encouraging in that both CON6 gp120 and gp140CF induced antibodies that neutralized select subtype

B, C and D HIV-1 primary isolates. However, it is clear that the most difficult-to-neutralize primary isolates (PAVO, 6101, BG1168, 92RW020, CM244) were either only weakly or not neutralized by anti-CON6 gp120 or gp140 sera (Table 1B). Nonetheless, the CON6 envelope immunogenicity for induction of neutralizing antibodies is promising, given the breadth of responses generated with a subunit envelope protein. Previous studies with poxvirus constructs expressing gp120 and gp160 have not generated high levels of neutralizing antibodies (Evans et al, J. Infect. Dis. 180:290-298 (1999), Polacino et al, J. Virol. 73:618-630 (1999), Ourmanov et al, J. Virol. 74:2960-2965 (2000), Pal et al, J. Virol 76:292-302 (2002), Excler and Plotkin, AIDS 11(Suppl A):S127-137 (1997). rVV expressing secreted CON6 gp120 and gp140 have been constructed and antibodies that neutralize HIV-1 primary isolates induced. It will be important to determine the breadth of neutralizing antibodies induced by DNA or recombinant adenovirus prime followed by recombinant vaccinia virus boost, to determine if expression of the CON6 envelope *in vivo* induces a broader neutralizing response than protein immunizations.

The structure of an oligomeric gp140 protein is critical when evaluating protein immunogenicity. In this regard, study of purified CON6 gp140CF proteins by fast performance liquid chromatography (FPLC) and analytical ultracentrifiguration has demonstrated

that the purified gp140 peak consists predominantly of trimers with a small component of dimers.

Thus, centralized envelopes such as CON6 are attractive candidates for preparation of various potentially "enhanced" envelope immunogens including CD4-Env complexes, constrained envelope structures, and trimeric oligomeric forms. The ability of CON6-induced T and B cell responses to protect against HIV-1 infection and/or disease in SHIV challenge models will be studied in non-human primates.

The above study has demonstrated that artificial centralized HIV-1 genes such as group M consensus *env* gene (CON6) can also induce T cell responses to T cell epitopes in wild-type subtype B and C Env proteins as well as to those on group M consensus Env proteins (Figure 5). While the DNA prime and rVV boost regimen with CON6 gp140CF immunogen clearly induced IFN-γ producing T cells that recognized subtype B and C epitopes, further studies are needed to determine if centralized sequences such as are found in the CON6 envelope are significantly better at inducing cross-clade T cell responses than wild-type HIV-1 genes (Ferrari et al, Proc. Natl. Acad. Sci. USA 94:1396-1401 (1997), Ferrari et al, AIDS Res. Hum. Retroviruses 16:1433-1443 (2000)). However, the fact that CON6 prime and boosted splenocyte T cells recognized HIV-1 subtype B and C T cell epitopes is an important step in demonstration that CON6 can induce T cell responses that might be clinically useful.

Three computer models (consensus, ancestor and center of the tree (COT)) have been proposed to generate centralized HIV-1 genes (Gaschen et al, Science 296:2354-2360 (2002), Gao et al, Science 299:1517-1518 (2003), Nickle et al, Science 299:1515-1517 (2003), Korber et al, Science 288:1789-1796 (2000). They all tend to locate at the roots of the star-like phylogenetic trees for most HIV-1 sequences within or between subtypes. As experimental vaccines, they all can reduce the genetic distances between immunogens and field virus strains. However, consensus, ancestral and COT sequences each have advantages and disadvantages (Gaschen et al, Science 296:2354-2360 (2002), Gao et al, Science 299:1517-1518 (2003), Nickle et al, Science 299:1515-1517 (2003). Consensus and COT represent the sequences or epitopes in sampled current wild-type viruses and are less affected by outliers HIV-1 sequences, while ancestor represents ancestral sequences that can be significantly affected by outlier sequences. However, at present, it is not known which centralized sequence can serve as the best immunogen to elicit broad immune responses against diverse HIV-1 strains, and studies are in progress to test these different strategies.

Taken together, the data have shown that the HIV-1 artificial CON6 envelope can induce T cell responses to wild-type HIV-1 epitopes, and can induce antibodies that neutralize HIV-1 primary isolates, thus demonstrating the feasibility and

promise of using artificial centralized HIV-1 sequences in HIV-1 vaccine design.

EXAMPLE 2

HIV-1 Subtype C Ancestral and Consensus Envelope Glycoproteins

EXPERIMENTAL DETAILS

HIV-1 subtype C ancestral and consensus *env* genes were obtained from the Los Alamos HIV Molecular Immunology Database (http://hiv-web.lanl.gov/immunology), codon-usage optimized for mammalian cell expression, and synthesized (Fig. 6). To ensure optimal expression, a Kozak sequence (GCCGCCGCC) was inserted immediately upstream of the initiation codon. In addition to the full-length genes, two truncated *env* genes were generated by introducing stop codons immediately after the gp41 membrane-spanning domain (IVNR) and the gp120/gp41 cleavage site (REKR), generating gp140 and gp120 form of the glycoproteins, respectively (Fig. 8).

Genes were tested for integrity in an *in vitro* transcription/translation system and expressed in mammalian cells. To determine if the ancestral and consensus subtype C envelopes were capable of mediating fusion and entry, *gp160* and *gp140* genes were co-transfected with an HIV-1/SG3Δ*env* provirus and the resulting pseudovirions tested for infectivity using the JC53-BL cell assay (Fig. 7). Co-receptor usage and envelope neutralization sensitivity were also determined with slight

modifications of the JC53-BL assay. Codon-usage optimized and rev-dependent 96ZAM651 *env* genes were used as contemporary subtype C controls.

RESULTS

Codon-optimized subtype C ancestral and consensus envelope genes (*gp160, gp140, gp120*) express high levels of *env* glycoprotein in mammalian cells (Fig. 9).

Codon-optimized subtype C *gp160* and *gp140* glycoproteins are efficiently incorporated into virus particles. Western Blot analysis of sucrose-purified pseudovirions reveals ten-fold higher levels of virion incorporation of the codon-optimized envelopes compared to that of a rev-dependent contemporary envelope controls (Fig. 10A).

Virions pseudotyped with either the subtype C consensus *gp160* or *gp140* envelope were more infectious than pseudovirions containing the corresponding gp160 and gp140 ancestral envelopes. Additionally, *gp160* envelopes were consistently more infectious than their respective *gp140* counterparts (Fig. 10B).

Both subtype C ancestral and consensus envelopes utilize CCR5 as a co-receptor to mediate virus entry (Fig. 11).

The infectivity of subtype C ancestral and consensus *gp160* containing pseudovirions was neutralized by plasma from subtype C infected patients. This suggests that these artificial

envelopes possess a structure that is similar to that of native HIV-1 *env* glycoproteins and that common neutralization epitopes are conserved. No significant differences in neutralization potential were noted between subtype C ancestral and consensus *env* glycoproteins (gp160) (Fig. 12).

CONCLUSIONS

HIV-1 subtype C viruses are among the most prevalent circulating isolates, representing approximately fifty percent of new infections worldwide. Genetic diversity among globally circulating HIV-1 strains poses a challenge for vaccine design. Although HIV-1 Env protein is highly variable, it can induce both humoral and cellular immune responses in the infected host. By analyzing 70 HIV-1 complete subtype C *env* sequences, consensus and ancestral subtype C *env* genes have been generated. Both sequences are roughly equidistant from contemporary subtype C strains and thus expected to induce better cross-protective immunity. A reconstructed ancestral or consensus sequence derived-immunogen minimizes the extent of genetic differences between the vaccine candidate and contemporary isolates. However, consensus and ancestral subtype C *env* genes differ by 5% amino acid sequences. Both consensus and ancestral sequences have been synthesized for analyses. Codon-optimized subtype C ancestral and consensus envelope genes have been constructed and the *in*

vitro biological properties of the expressed glycoproteins determined. Synthetic subtype C consensus and ancestral *env* genes express glycoproteins that are similar in their structure, function and antigenicity to contemporary subtype C wild-type envelope glycoproteins.

EXAMPLE 3

Codon-Usage Optimization of Consensus of Subtype C *gag* and *nef* Genes (C.con.gag and C.con.nef)

Subtype C viruses have become the most prevalent viruses among all subtypes of Group M viruses in the world. More than 50% of HIV-1 infected people are currently carrying HIV-1 subtype C viruses. In addition, there is considerable intra-subtype C variability: different subtype C viruses can differ by as much as 10%, 6%, 17% and 16% of their Gag, Pol, Env and Nef proteins, respectively. Most importantly, the subtype C viruses from one country can vary as much as the viruses isolated from other parts of the world. The only exceptions are HIV-1 strains from India/China, Brazil and Ethiopia/Djibouti where subtype C appears to have been introduced more recently. Due to the high genetic variability of subtype C viruses even within a single country, an immunogen based on a single virus isolate may not elicit protective

immunity against other isolates circulating in the same area.

Thus *gag* and *nef* gene sequences of subtype C viruses were gathered to generate consensus sequences for both genes by using a 50% consensus threshold. To avoid a potential bias toward founder viruses, only one sequence was used from India/China, Brazil and Ethiopia/Djibouti, respectively, to generate the subtype C consensus sequences (C.con.gag and C.con.nef). The codons of both C.con.gag and C.con.nef genes were optimized based on the codon usage of highly expressed human genes. The protein expression following transfection into 293T cells is shown in Figure 13. As can be seen, both consensus subtype C Gag and Nef proteins were expressed efficiently and recognized by Gag- and Nef-specific antibodies. The protein expression levels of both C.con.gag and C.con.nef genes are comparible to that of native subtype *env* gene (96ZM651).

EXAMPLE 4

Synthesis of a Full Length "*Consensus of the Consensus env* Gene with Consensus Variable Regions" (CONs)

In the synthesized "consensus of the consensus" *env* gene (CON6), the variable regions were replaced

with the corresponding regions from a contemporary subtype C virus (98CN006). A further con/con gene has been designed that also has consensus variable regions (CONs). The codons of the CONs *env* gene were optimized based on the codon usage of highly expressed human genes. (See Figs. 14A and 14B for amino acid sequences and nucleic acid sequences, respectfully.)

Paired oligonucleotides (80-mers) which overlap by 20 bp at their 3' ends and contain invariant sequences at their 5' and 3' ends, including the restriction enzyme sites EcoRI and BbsI as well as BsmBI and BamHI, respectively, were designed. BbsI and BamHI are Type II restriction enzymes that cleave outside of their recognition sequences. They have been positioned in the oligomers in such a way that they cleave the first four resides adjacent to the 18 bp invariant region, leaving 4 base 5' overhangs at the end of each fragment for the following ligation step. 26 paired oligomers were linked individually using PCR and primers complimentary to the 18 bp invariant sequences. Each pair was cloned into pGEM-T (Promega) using the T/A cloning method and sequenced to confirm the absence of inadvertent mutations/deletions. pGEM-T subclones containing the proper inserts were then digested, run on a 1% agarose gel, and gel purified (Qiagen). Four individual 108-mers were ligated into pcDNA3.1 (Invitrogen) in a multi-fragment ligation reaction. The four-way ligations occurred among groups of fragments in a stepwise manner from

the 5′ to the 3′ end of the gene. This process was repeated until the entire gene was reconstructed in the pcDNA3.1 vector.

A complete CONs gene was constructed by ligating the codon usage optimized oligo pairs together. To confirm its open reading frame, an *in vitro* transcription and translation assay was performed. Protein products were labeled by S^{35}-methionine during the translation step, separated on a 10% SDS-PAGE, and detected by radioautography. Expected size of the expressed CONS gp160 was identified in 4 out of 7 clones (Fig. 14C).

CONs Env protein expression in the mammalian cells after transfected into 293T cells using a Western blot assay (Figure 15). The expression level of CONs Env protein is very similar to what was observed from the previous CON6 *env* clone that contains the consensus conservative regions and variable loops from 98CN006 virus isolate.

The Env-pseudovirons was produced by cotransfecting CONs *env* clone and env-deficient SG3 proviral clone into 293T cells. Two days after transfection, the pseudovirions were harvested and infected into JC53BL-13 cells. The infectious units (IU) were determined by counting the blue cells after staining with X-gal in three independent experiments. When compared with CON6 *env* clone, CONs *env* clones produce similar number of IU in JC53BL-13 cells (Figure 16). The IU titers for both are about 3 log higher than the SG3 backbone clone control (No

Env). However, the titers are also about 2 log lower than the positive control (the native HIV-1 *env* gene, NL4-3 or YU2). These data suggest that both *consensus* group M *env* clones are biologically functional. Their functionality, however, has been compromised. The functional consensus *env* genes indicate that these Env proteins fold correctly, preserve the basic conformation of the native Env proteins, and are able to be developed as universal Env immunogens.

It was next determined what coreceptor CONs Env uses for its entry into JC53-BL cells. When treated with CXCR4 blocking agent AMD3100, the infectivity of NL4-3 Env-pseudovirons was blocked while the infectivity of YU2, CONs or CON6 Env-pseudovirons was not inhibited. In contrast, when treated with CCR5 blocking agent TAK779, the infectivity of NL4-3 Env-pseudovirons was not affected, while the infectivity of YU2, CONs or CON6 Env-pseudovirons was inhibited. When treated with both blocking agents, the infectivity of all pseudovirions was inhibited. Taken together, these data show that the CONs as well as CON6 envelope uses the CCR5 but not CXCR4 co-receptor for its entry into target cells.

It was next determined whether CON6 or CONs Env proteins could be equally efficiently incorporated in to the pseudovirions. To be able precisely compare how much Env proteins were incorporated into the pseudovirions, each pseudovirions is loaded on SDS-PAGE at the same concentraion: 5μg total protein

for cell lysate, 25ng p24 for cell culture supernatant, or 150ng p24 for purified virus stock (concentrated pseudovirions after super-speed centrifugation). There was no difference in amounts of Env proteins incorporated in CON6 or CONs Env-pseudovirions in any preparations (cell lysate, cell culture supernatant or purified virus stock) (Figure 17).

EXAMPLE 5

Synthesis of a *Consensus* Subtype A Full Length *env* (A.con.env) Gene

Subtype A viruses are the second most prevalent HIV-1 in the African continent where over 70% of HIV-1 infections have been documented. Consensus *gag*, *env* and *nef* genes for subtype C viruses that are the most prevalent viruses in Africa and in the world were previously generated. Since genetic distances between subtype A and C viruses are as high as 30% in the *env* gene, the cross reactivity or protection between both subtypes will not be optimal. Two group M consensus *env* genes for all subtypes were also generated. However, to target any particular subtype viruses, the subtype specific consensus genes will be more effective since the genetic distances between subtype consensus genes and field viruses from the same subtype will be smaller than that between group M consensus genes and these same viruses. Therefore, consensus genes need to be generated for development of subtype A

specific immunogens. The codons of the A.con.env gene were optimized based on the codon usage of highly expressed human genes. (See Figs. 18A and 18B for amino acid and nucleic acid sequences, respectively.)

Each pair of the oligos has been amplified, cloned, ligated and sequenced. After the open reading frame of the A.con *env* gene was confirmed by an *in vitro* transcription and translation system, the A.con *env* gene was transfected into the 293T cells and the protein expression and specificity confirmed with the Western blot assay (Figure 18). It was then determined whether A.con envelope is biologically functional. It was co-transfected with the *env*-defective SG3 proviral clone into 293T cells. The pseudotyped viruses were harvested and used to infect JC53BL cells. Blue cells were detected in JC53-BL cells infected with the A.con Env-pseudovirions, suggesting that A.con Env protein is biologically functional (Table 2). However, the infectious titer of A.con Env-psuedovirions was about 7-fold lower than that of pseudovirions with wild-type subtype C envelope (Table 2). Taken together, the biological function A.con Env proteins suggests that it folds correctly and may induce linear and conformational T and B cell epitopes if used as an Env immunogen.

JC53BL13 (IU/ul)

	3/31/03	4/7/03	4/25/03
	non filtered supt.	0.22µm filtered	0.22µm filtered
A.con +SG3	4	8.5	15.3
96ZM651 +SG3	87	133	104
SG3 backbone	0	0.07	0.03
Neg control	0	0.007	0

Table 2. Infectivity of pseudovirons with A.con env genes

EXAMPLE 6

Design of Full Length "*Consensus of the Consensus gag, pol and nef* Genes" (M.con.gag, M.con.pol and M.con.nef) and a Subtype C *Consensus* pol Gene (C.con.pol)

For the group M consensus genes, two different *env* genes were constructed, one with virus specific variable regions (CON6) and one with consensus variable regions (CONs). However, analysis of T cell immune responses in immunized or vaccinated animals and humans shows that the *env* gene normally is not a main target for T cell immune response although it is the only gene that will induce neutralizing antibody. Instead, HIV-1 Gag, Pol and Nef proteins are found to be important for inducing potent T cell immune responses. To generate a repertoire of immunogens that can induce both broader humoral and cellular immune responses for

all subtypes, it may be necessary to construct other group M consensus genes other than *env* gene alone. "*Consensus of the consensus*" *gag*, *pol* and *nef* genes (M.con.gag., M.con.pol and M.con.nef) have been designed. To generate a subtype consensus pol gene, the subtype C consensus *pol* gene (C.con.pol) was also designed. The codons of the M.con.gag., M.con.pol, M.con.nef and C.con.pol. genes were optimized based on the codon usage of highly expressed human genes. (See Fig. 19 for nucleic acid and amino acid sequences.)

EXAMPLE 7

Synthetic Subtype B Consensus *gag* and *env* Genes

EXPERIMENTAL DETAILS

Subtype B consensus *gag* and *env* sequences were derived from 37 and 137 contemporary HIV-1 strains, respectively, codon-usage optimized for mammalian cell expression, and synthesized (Figs. 20A and 20B). To ensure optimal expression, a Kozak sequence (GCCGCCGCC) was inserted immediately upstream of the initiation codon. In addition to the full-length *env* gene, a truncated *env* gene was generated by introducing a stop codon immediately after the gp41 membrane-spanning domain (IVNR) to create a *gp145* gene. Genes were tested for integrity in an *in vitro* transcription/translation system and expressed in mammalian cells. (Subtype B consensus Gag and Env sequences are set forth in Figs. 20C and 20D, respectively.)

To determine if the subtype B consensus envelopes were capable of mediating fusion and entry, *gp160* and *gp145* genes were co-transfected with an HIV-1/SG3Δenv provirus and the resulting pseudovirions were tested for infectivity using the JC53-BL cell assay. JC53-BL cells are a derivative of HeLa cells that express high levels of CD4 and the HIV-1 coreceptors CCR5 and CXCR4. They also contain the reporter cassettes of luciferase and β-galactosidase that are each expressed from an HIV-1 LTR. Expression of the reporter genes is dependent on production of HIV-1 Tat. Briefly, cells are seeded into 24-well plates, incubated at 37°C for 24 hours and treated with DEAE-Dextran at 37°C for 30min. Virus is serially diluted in 1% DMEM, added to the cells incubating in DEAE-dextran, and allowed to incubate for 3 hours at 37°C after which an additional 500μL of cell media is added to each well. Following a final 48-hour incubation at 37°C, cells are fixed, stained using X-Gal, and overlaid with PBS for microscopic counting of blue foci. Counts for mock-infected wells, used to determine background, are subtracted from counts for the sample wells. Co-receptor usage and envelope neutralization sensitivity were also determined with slight modifications of the JC53-BL assay.

To determine whether the subtype B consensus Gag protein was capable of producing virus-like particles (VLPs) that incorporated Env glycoproteins, 293T cells were co-transfected with

subtype B consensus *gag* and *env* genes. 48-hours post-transfection, cell supernatants containing VLPs were collected, clarified in a tabletop centrifuge, filtered through a 0.2mM filter, and pellet through a 20% sucrose cushion. The VLP pellet was resuspended in PBS and transferred onto a 20-60% continuous sucrose gradient. Following overnight centrifugation at 100,000 x g, 0.5 ml fractions were collected and assayed for p24 content. The refractive index of each fraction was also measured. Fractions with the correct density for VLPs and containing the highest levels of p24 were pooled and pellet a final time. VLP-containing pellets were re-suspended in PBS and loaded on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with serum from a subtype B HIV-1 infected individual.

RESULTS

Codon-usage optimized, subtype B consensus envelope (*gp160, gp145*) and *gag* genes express high levels of glycoprotein in mammalian cells (Fig. 21).

Subtype B gp160 and gp145 glycoproteins are efficiently incorporated into virus particles. Western Blot analysis of sucrose-purified pseudovirions suggests at least five-fold higher levels of consensus B envelope incorporation compared to incorporation of a rev-dependent contemporary envelope (Fig.23A). Virions pseudotyped with either the subtype B consensus

gp160 or gp145 envelope are more infectious than pseudovirions containing a rev-dependent contemporary envelope (Fig. 23 B).

Subtype B consensus envelopes utilize CCR5 as the co-receptor to gain entry into CD4 bearing target cells (Fig. 22).

The infectivity of pseudovirions containing the subtype B consensus gp160 envelope was neutralized by plasma from HIV-1 subtype B infected patients (Fig. 24C) and neutralizing monoclonal antibodies (Fig. 24A). This suggests that the subtype B synthetic consensus B envelopes is similar to native HIV-1 Env glycoproteins in its overall structure and that common neutralization epitopes remain intact. Figs. 24B and 24D show neutralization profiles of a subtype B control envelope (NL4.3 Env).

Subtype B consensus Gag proteins are able to bud from the cell membrane and form virus-like particles (Fig. 25A). Co-transfection of the codon-optimized subtype B consensus *gag* and *gp160* genes produces VLPs with incorporated envelope (Fig. 25B).

CONCLUSIONS

The synthetic subtype B consensus *env* and *gag* genes express viral proteins that are similar in their structure, function and antigenicity to contemporary subtype B Env and Gag proteins. It is contemplated that immunogens based on subtype B consensus genes will elicit CTL and neutralizing

immune responses that are protective against a broad set of HIV-1 isolates.

* * *

All documents and other information sources cited above are hereby incorporated in their entirety by reference.

ABSTRACT

The present invention relates, in general, to an immunogen and, in particular, to an immunogen for inducing antibodies that neutralize a wide spectrum of HIV primary isolates and/or to an immunogen that induces a T cell immune response. The invention also relates to a method of inducing anti-HIV antibodies, and/or to a method of inducing a T cell immune response, using such an immunogen. The invention further relates to nucleic acid sequences encoding the present immunogens.

A

MRVMGIQRNCQHLWRWGTMILGMLMICSAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNVWAT

V1

HACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNVRNVSSNG

V2

TETDNEEIKNCSFNITTELRDKKQKVYALFYRLDVVPIDDKNSSEISGKNSSEYYRLINCNTSAITQACP

KVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSEN

V3

ITNNAKTIIVQLNESVEINCTRPNNNTRKSIHIGPGQAFYATGEIIGDIRQAHCNISRTKWNKTLQQVAK

V4

KLREHFNNKTIIFKPSSGGDLEITTHSFNCGGEFFYCNTSGLFNSTWMFNGTYMFNGTKDNSETITLPCR

V5

IKQIINMWQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNSNKNKTETFRPGGGDMRDNWRSELYKYK

VVKIEPLGVAPTKAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLR

AIEAQQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTNVPWNSSWSNKSQDEIWDNMT

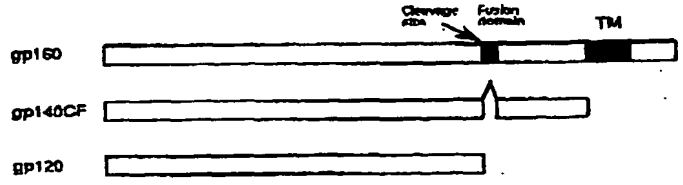
WMEWEREISNYTDIIYRLIEESQNQQEKNEQELLALDKWASLWNWFDITNWLWYIKIFIMIVGGLIGLRI

VFAVLSIVNRVRQGYSPLSFQTLIPNPRGPDRPEGIEEEGGEQGRDRSIRLVNGFLALAWDDLRSLCLFS

YHRLRDFILIAARTVELLGRRSLRGLQKGWEALKYLGNLLQYWGQELKNSAISLLDTTAIAVAEGTDRVI

EIVQRACRAILNIPRRIRQGLERALL

B



C

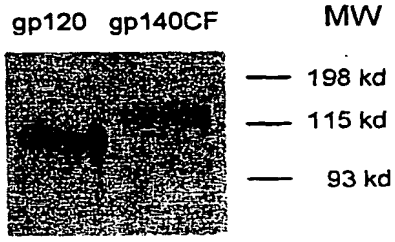


Figure 1

CON6.env (group M env consensus. This one contain five variable regions in env gene from 98CN006 virus, not in the public domain yet)

GCCACCATGCGCGTGATGGGCATCCAGCGCAACTGCCAGCACCTGTGGCGCTGGGGCACCATGATC
CTGGGCATGCTGATGATCTGCTCCGCCGCCGAGAACCTGTGGGTGACCGTGTACTACGGC
GTGCCCGTGTGGAAGGAGGCCAACACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTAC
GACACCGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCC
CAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTG
GAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAG
CTGACCCCCCTGTGCGTGACCCTGAACTGCACCAACGTGCGCAACGTGTCCTCCAACGGC
ACCGAGACCGACAACGAGGAGATCAAGAACTGCTCCTTCAACATCACCACCGAGCTGCGC
GACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCGACGAC
AAGAACTCCTCCGAGATCTCCGGCAAGAACTCCTCCGAGTACTACCGCCTGATCAACTGC
AACACCTCCGCCATCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCCAC
TACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACC
GGCCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCC
ACCCAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAAC
ATCACCAACAACGCCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGC
ACCCGCCCCAACAACAACACCCGCAAGTCCATCCACATCGGCCCCGGCCAGGCCTTCTAC
GCCACCGGCGAGATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCCGCACCAAG
TGGAACAAGACCCTGCAGCAGGTGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACC
ATCATCTTCAAGCCCTCCTCCGGCGGCGACCTGGAGATCACCACCCACTCCTTCAACTGC
GGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCCACCTGGATGTTCAAC
GGCACCTACATGTTCAACGGCACCAAGGACAACTCCGAGACCATCACCCTGCCCTGCCGC
ATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATC
GAGGGCAAGATCACCTGCAAGTCCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGC
AACAACTCCAACAAGAACAAGACCGAGACCTTCCGCCCCGGCGGCGGCGACATGCGCGAC
AACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCC
CCCACCAAGGCCAAGCGCCGCGTGGTGGAGCGCGAGAAGCGCGCCGTGGGCATCGGCGCC
GTGTTCCTGGGCTTCCTGGGCGCCGCCGGCTCCACCATGGGCGCCGCCTCCATCACCCTG
ACCGTGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCCAACCTGCTGCGC
GCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAG
GCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGC
TGCTCCGGCAAGCTGATCTGCACCACCAACGTGCCCTGGAACTCCTCCTGGTCCAACAAG
TCCCAGGACGAGATCTGGGACAACATGACCTGGATGGAGTGGGAGCGCGAGATCTCCAAC
TACACCGACATCATCTACCGCCTGATCGAGGAGTCCCAGAACCAGCAGGAGAAGAACGAG
CAGGAGCTGCTGGCCCTGGACAAGTGGGCCTCCCTGTGGAACTGGTTCGACATCACCAAC
TGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGATCGGCCTGCGCATC
GTGTTCGCCGTGCTGTCCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCCCCTGTCCTTC
CAGACCCTGATCCCCAACCCCCGCGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGC
GGCGAGCAGGGCCGCGACCGCTCCATCCGCCTGGTGAACGGCTTCCTGGCCCTGGCCTGG
GACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCGCCTGCGCGACTTCATCCTGATC
GCCGCCCGCACCGTGGAGCTGCTGGGCCGCCGCTCCCTGCGCGGCCTGCAGAAGGGCTGG
GAGGCCCTGAAGTACCTGGGCAACCTGCTGCAGTACTGGGGCCAGGAGCTGAAGAACTCC
GCCATCTCCCTGCTGGACACCACCGCCATCGCCGTGGCCGAGGGCACCGACCGCGTGATC
GAGATCGTGCAGCGCGCCTGCCGCGCCATCCTGAACATCCCCCGCCGCATCCGCCAGGGC
CTGGAGCGCGCCCTGCTGTAA

Figure D

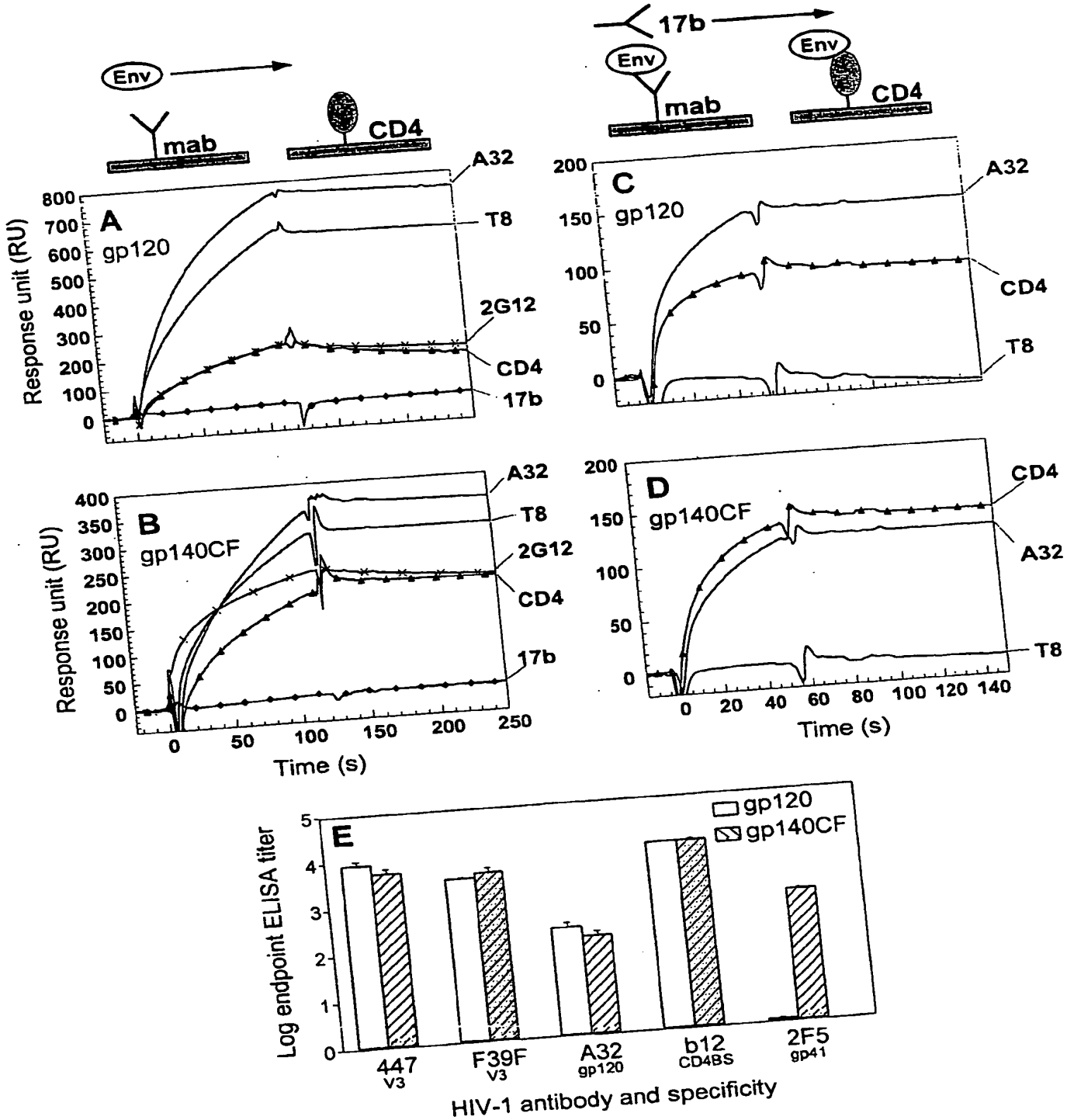

Env
mab
CD4
A
gp120
Response unit (RU)
A32
T8
2G12
CD4
17b
B
gp140CF
Time (s)
17b
Env
C
D
E
gp120
gp140CF
Log endpoint ELISA titer
447
V3
F39F
V3
A32
gp120
b12
CD4BS
2F5
gp41
HIV-1 antibody and specificity

Figure 2

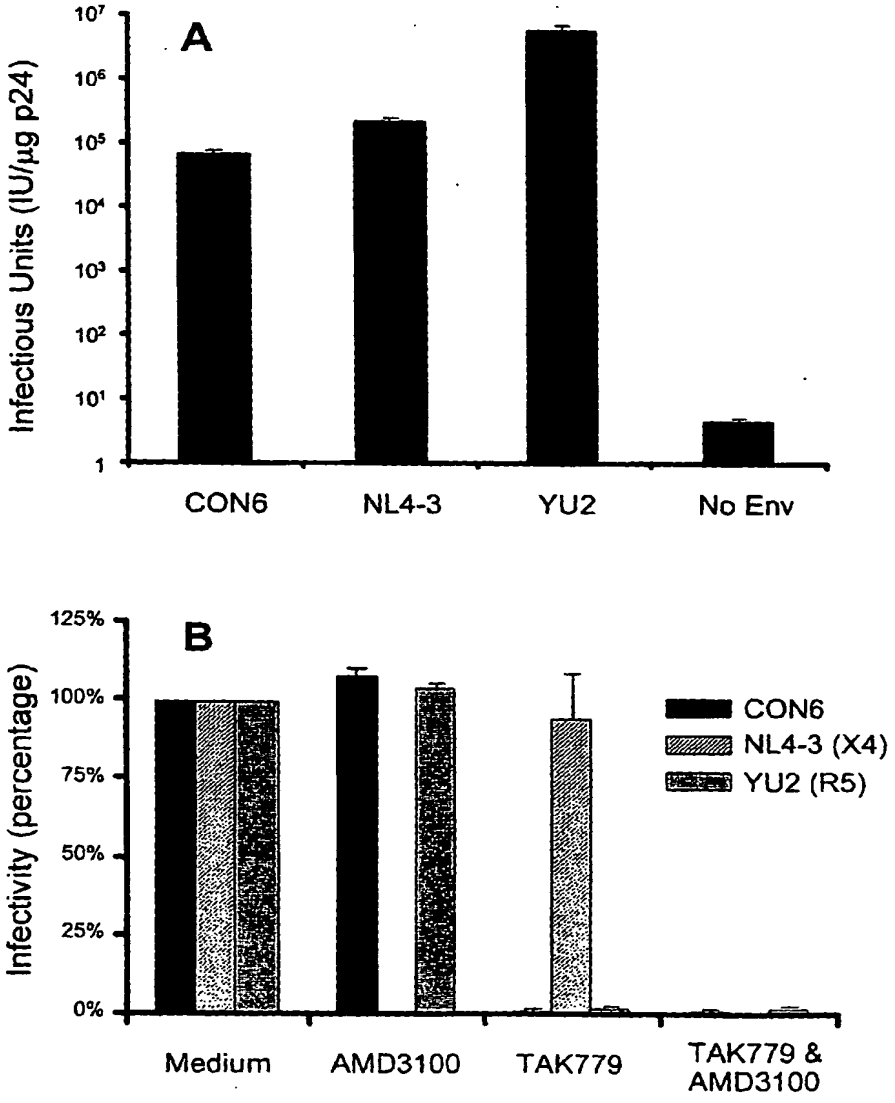

A
Infectious Units (IU/µg p24)
10⁷
10⁶
10⁵
10⁴
10³
10²
10¹
1
CON6
NL4-3
YU2
No Env
B
Infectivity (percentage)
125%
100%
75%
50%
25%
0%
CON6
NL4-3 (X4)
YU2 (R5)
Medium
AMD3100
TAK779
TAK779 & AMD3100

Figure 3

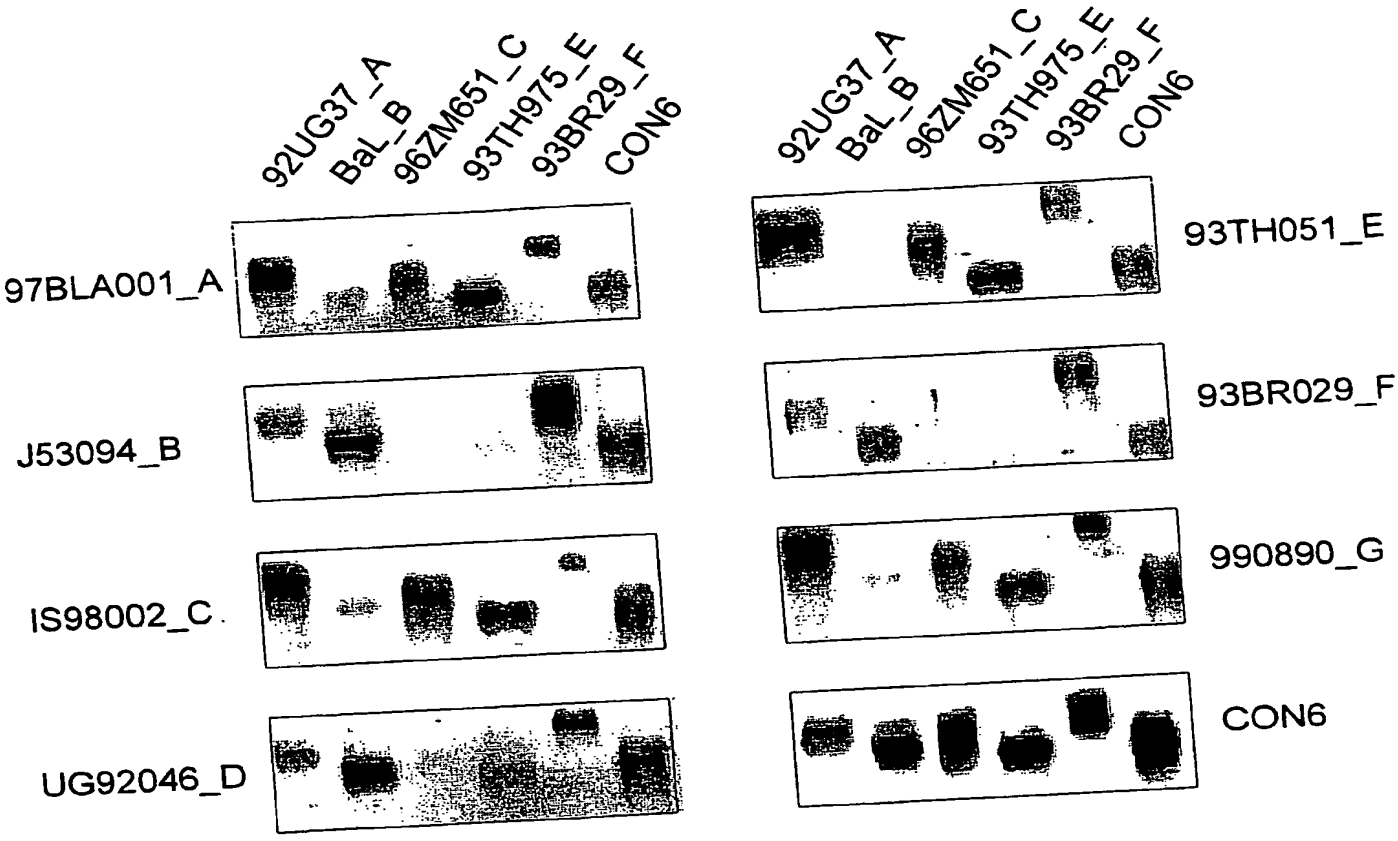

92UG37_A
BaL_B
96ZM651_C
93TH975_E
93BR29_F
CON6
97BLA001_A
J53094_B
IS98002_C
UG92046_D
93TH051_E
93BR029_F
990890_G
CON6

Figure 4

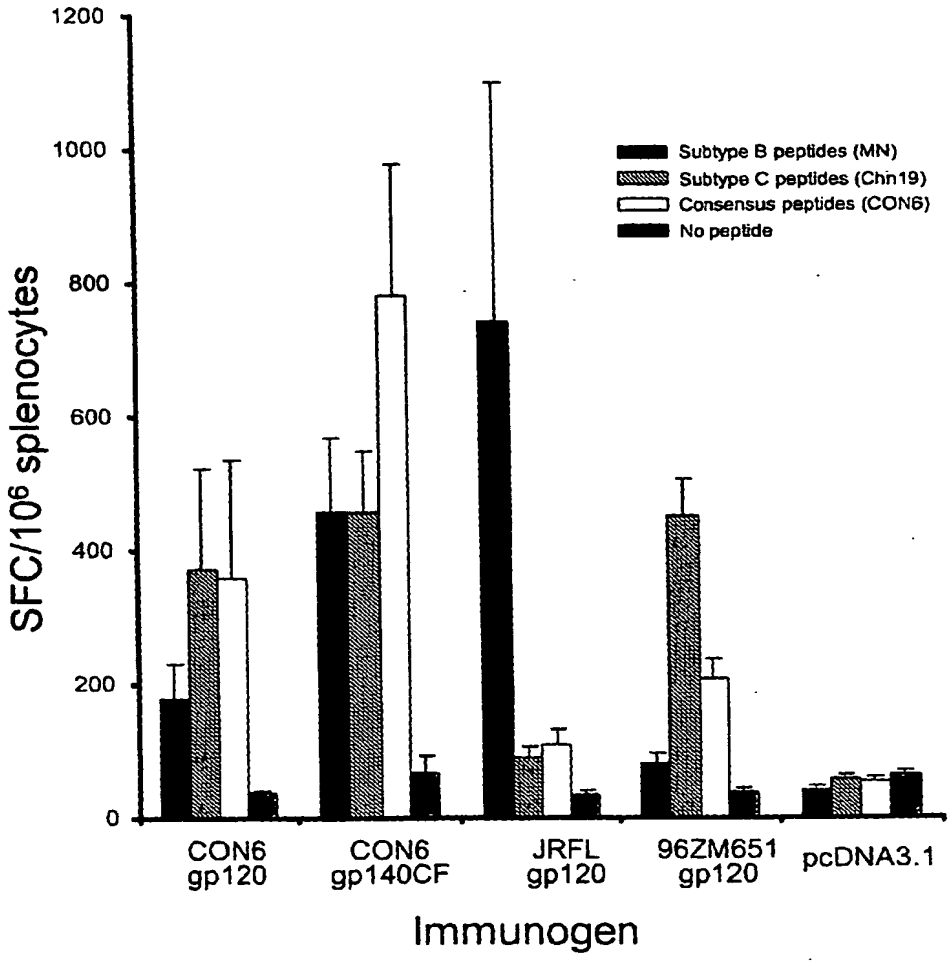

1200
1000
800
600
400
200
0
SFC/10^6 splenocytes
Subtype B peptides (MN)
Subtype C peptides (Chn19)
Consensus peptides (CON6)
No peptide
CON6 gp120
CON6 gp140CF
JRFL gp120
96ZM651 gp120
pcDNA3.1
Immunogen

Figure 5

Figure 6A

C.anc.env (subtype C ancestral env. The amino acid sequence is different from Los Alamos Database August 2002)

GCCGCCATGCGCGTGATGGGCATCCTGCGCAACTGCCAGCAGTGGTGGAT
CTGGGGCATCCTGGGCTTCTGGATGCTGATGATCTGCTCCGTGGTGGGCA
ACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAAG
ACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGAGCGCGAGGTGCA
CAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGG
AGATGGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACGAC
ATGGTGGACCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCT
GAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCACCA
ACGTGACCAACGCCACCAACAACACCTACAACGGCGAGATGAAGAACTGC
TCCTTCAACATCACCACCGAGCTGCGCGACAAGAAGAAGAAGGAGTACGC
CCTGTTCTACCGCCTGGACATCGTGCCCCTGAACGAGAACTCCTCCGAGT
ACCGCCTGATCAACTGCAACACCTCCGCCATCACCCAGGCCTGCCCCAAG
GTGTCCTTCGACCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTACGC
CATCCTGAAGTGCAACAACAAGACCTTCAACGGCACCGGCCCCTGCAACA
ACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACC
CAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTC
CGAGAACCTGACCGACAACGCCAAGACCATCATCGTGCAGCTGAACGAGT
CCGTGGAGATCGTGTGCACCCGCCCCAACAACAACACCCGCAAGTCCATG
CGCATCGGCCCCGGCCAGACCTTCTACGCCACCGGCGACATCATCGGCGA
CATCCGCCAGGCCCACTGCAACATCTCCGAGGACAAGTGGAACAAGACCC
TGCAGCAGGTGGCCGAGAAGCTGGGCAAGCACTTCCCCAACAAGACCATC
ACCTTCGAGCCCTCCTCCGGCGGCGACCTGGAGATCACCACCCACTCCTT
CAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCAAGCTGTTCAACT
CCACCTACAACAACAACACCAACTCCAACTCCACCATCACCCTGCCCTGC
CGCATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTA
CGCCCCCCCCATCGCCGGCAACATCACCTGCAAGTCCAACATCACCGGCC
TGCTGCTGACCCGCGACGGCGGCAAGGAGAACACCACCGAGACCTTCCGC
CCCGGCGGCGGCGACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTA
CAAGGTGGTGGAGATCAAGCCCCTGGGCGTGGCCCCCACCGAGGCCAAGC
GCCGCGTGGTGGAGCGCGAGAAGCGCGCCGTGGGCCTGGGCGCCGTGTTC
CTGGGCTTCCTGGGCGCCGCCGGCTCCACCATGGGCGCCGCCTCCATCAC
CCTGACCGTGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGT
CCAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACATGCTGCAGCTGACC
GTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCATGGAGCGCTA
CCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGA
TCTGCACCACCGCCGTGCCCTGGAACTCCTCCTGGTCCAACAAGTCCCTG
GACGACATCTGGGACAACATGACCTGGATGGAGTGGGACCGCGAGATCTC
CAACTACACCGACACCATCTACCGCCTGCTGGAGGAGTCCCAGAACCAGC
AGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACTCCTGGGAGAACCTG
TGGAACTGGTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCAT
CATGATCGTGGGCGGCCTGATCGGCCTGCGCATCATCTTCGCCGTGCTGT
CCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCCCTGTCCTTCCAGACC
CTGACCCCCAACCCCCGCGGCCCCGACCGCCTGGAGCGCATCGAGGAGGA
GGGCGGCGAGCAGGACCGCGACCGCTCCATCCGCCTGGTGTCCGGCTTCC
TGGCCCTGGCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCAC
CGCCTGCGCGACTTCATCCTGATCGCCGCCCGCACCGTGGAGCTGCTGGG
CCGCTCCTCCCTGCGCGGCCTGCAGCGCGGCTGGGAGGCCCTGAAGTACC
TGGGCTCCCTGGTGCAGTACTGGGGCCAGGAGCTGAAGAAGTCCGCCATC
TCCCTGCTGGACACCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCAT
CATCGAGGTGGTGCAGCGCGCCTGCCGCGCCATCCTGAACATCCCCCGCC
GCATCCGCCAGGGCTTCGAGGCCGCCCTGCTGTAA

Figure 6B

C.con.env (subtype C consensus env. The amino acid sequence is different from Los Alamos Database August 2002)

GCCGCCATGCGCGTGATGGGCATCCTGCGCAACTGCCAGCAGTGGTGGAT
CTGGGGCATCCTGGGCTTCTGGATGCTGATGATCTGCAACGTGGTGGGCA
ACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAAG
ACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGAGAAGGAGGTGCA
CAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGG
AGATGGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACGAC
ATGGTGGACCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCT
GAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCCGCA
ACGTGACCAACGCCACCAACAACACCTACAACGAGGAGATCAAGAACTGC
TCCTTCAACATCACCACCGAGCTGCGCGACAAGAAGAAGAAGGTGTACGC
CCTGTTCTACCGCCTGGACATCGTGCCCCTGAACGAGAACTCCTCCGAGT
ACCGCCTGATCAACTGCAACACCTCCGCCATCACCCAGGCCTGCCCCAAG
GTGTCCTTCGACCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTACGC
CATCCTGAAGTGCAACAACAAGACCTTCAACGGCACCGGCCCCTGCAACA
ACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACC
CAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTC
CGAGAACCTGACCAACAACGCCAAGACCATCATCGTGCACCTGAACGAGT
CCGTGGAGATCGTGTGCACCCGCCCCAACAACAACACCCGCAAGTCCATC
CGCATCGGCCCCGGCCAGACCTTCTACGCCACCGGCGACATCATCGGCGA
CATCCGCCAGGCCCACTGCAACATCTCCGAGGACAAGTGGAACAAGACCC
TGCAGCGCGTGTCCAAGAAGCTGAAGGAGCACTTCCCCAACAAGACCATC
AAGTTCGAGCCCTCCTCCGGCGGCGACCTGGAGATCACCACCCACTCCTT
CAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCAAGCTGTTCAACT
CCACCTACAACAACAACACCAACTCCAACTCCACCATCACCCTGCCCTGC
CGCATCAAGCAGATCATCAACATGTGGCAGGAGGTGGGCCGCGCCATGTA
CGCCCCCCCCATCGCCGGCAACATCACCTGCAAGTCCAACATCACCGGCC
TGCTGCTGACCCGCGACGGCGGCAAGAAGAACACCACCGAGATCTTCCGC
CCCGGCGGCGGCGACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTA
CAAGGTGGTGGAGATCAAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGC
GCCGCGTGGTGGAGCGCGAGAAGCGCGCCGTGGGCATCGGCGCCGTGTTC
CTGGGCTTCCTGGGCGCCGCCGGCTCCACCATGGGCGCCGCCTCCATCAC
CCTGACCGTGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGT
CCAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACATGCTGCAGCTGACC
GTGTGGGGCATCAAGCAGCTGCAGACCCGCGTGCTGGCCATCGAGCGCTA
CCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGA
TCTGCACCACCGCCGTGCCCTGGAACTCCTCCTGGTCCAACAAGTCCCAG
GAGGACATCTGGGACAACATGACCTGGATGCAGTGGGACCGCGAGATCTC
CAACTACACCGACACCATCTACCGCCTGCTGGAGGACTCCCAGAACCAGC
AGGAGAAGAACGAGAAGGACCTGCTGGCCCTGGACTCCTGGAAGAACCTG
TGGAACTGGTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCAT
CATGATCGTGGGCGGCCTGATCGGCCTGCGCATCATCTTCGCCGTGCTGT
CCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCCCTGTCCTTCCAGACC
CTGACCCCCAACCCCCGCGGCCCCGACCGCCTGGGCCGCATCGAGGAGGA
GGGCGGCGAGCAGGACCGCGACCGCTCCATCCGCCTGGTGTCCGGCTTCC
TGGCCCTGGCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCAC
CGCCTGCGCGACTTCATCCTGGTGGCCGCCCGCGCCGTGGAGCTGCTGGG
CCGCTCCTCCCTGCGCGGCCTGCAGCGCGGCTGGGAGGCCCTGAAGTACC
TGGGCTCCCTGGTGCAGTACTGGGGCCTGGAGCTGAAGAAGTCCGCCATC
TCCCTGCTGGACACCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCAT
CATCGAGCTGATCCAGCGCATCTGCCGCGCCATCCGCAACATCCCCCGCC
GCATCCGCCAGGGCTTCGAGGCCGCCCTGCAGTAA

Figure 6C

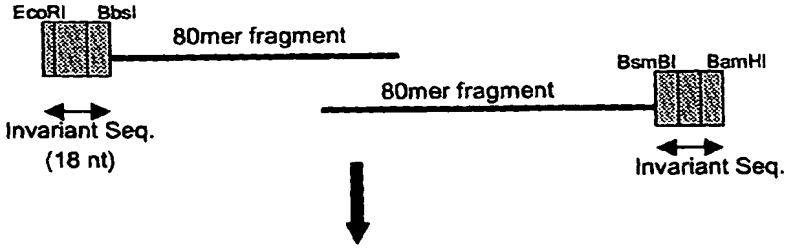
C.anc.env (subtype C ancestral env)
MRVMGILRNCQQWWIWGILGFWMLMICSVVGNLWVTVYYGVPVWKEAKTTLFCASDAKAYEREVHNVWAT
HACVPTDPNPQEMVLENVTENFNMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNVTNATNNT
YNGEMKNCSFNITTELRDKKKKEYALFYRLDIVPLNENSSEYRLINCNTSAITQACPKVSFDPIPIHYCA
PAGYAILKCNNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNAKTIIVQLN
ESVEIVCTRPNNNTRKSMRIGPGQTFYATGDIIGDIRQAHCNISEDKWNKTLQQVAEKLGKHFPNKTITF
EPSSGGDLEITTHSFNCRGEFFYCNTSKLFNSTYNNNTNSNSTITLPCRIKQIINMWQGVGQAMYAPPIA
GNITCKSNITGLLLTRDGGKENTTETFRPGGGDMRDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKR
AVGLGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVL
AMERYLKDQQLLGIWGCSGKLICTTAVPWNSSWSNKSLDDIWDNMTWMEWDREISNYTDTIYRLLEESQN
QQEKNEQDLLALDSWENLWNWFDITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTLT
PNPRGPDRLERIEEEGGEQDRDRSIRLVSGFLALAWDDLRSLCLFSYHRLRDFILIAARTVELLGRSSLR
GLQRGWEALKYLGSLVQYWGQELKKSAISLLDTIAIAVAEGTDRIIEVVQRACRAILNIPRRIRQGFEAA
LL

Figure 6D

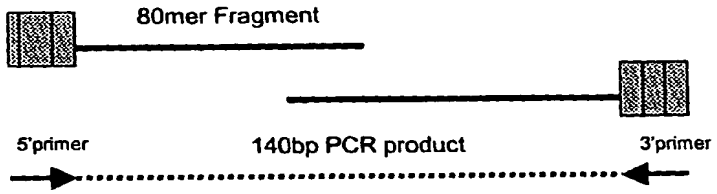
C.con.env (subtype C consensus env)
MRVMGILRNCQQWWIWGILGFWMLMICNVVGNLWVTVYYGVPVWKEAKTTLFCASDAKAYEKEVHNVWAT
HACVPTDPNPQEMVLENVTENFNMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCRNVTNATNNT
YNEEIKNCSFNITTELRDKKKKVYALFYRLDIVPLNENSSEYRLINCNTSAITQACPKVSFDPIPIHYCA
PAGYAILKCNNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTNNAKTIIVHLN
ESVEIVCTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAHCNISEDKWNKTLQRVSKKLKEHFPNKTIKF
EPSSGGDLEITTHSFNCRGEFFYCNTSKLFNSTYNNNTNSNSTITLPCRIKQIINMWQEVGRAMYAPPIA
GNITCKSNITGLLLTRDGGKKNTTEIFRPGGGDMRDNWRSELYKYKVVEIKPLGVAPTKAKRRVVEREKR
AVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQTRVL
AIERYLKDQQLLGIWGCSGKLICTTAVPWNSSWSNKSQEDIWDNMTWMQWDREISNYTDTIYRLLEDSQN
QQEKNEKDLLALDSWKNLWNWFDITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTLT
PNPRGPDRLGRIEEEGGEQDRDRSIRLVSGFLALAWDDLRSLCLFSYHRLRDFILVAARAVELLGRSSLR
GLQRGWEALKYLGSLVQYWGLELKKSAISLLDTIAIAVAEGTDRIIELIQRICRAIRNIPRRIRQGFEAA
LQ

Figure 6E

Synthesize entire gene in 80-mer fragments overlapping by 20 residues at the 3' end with invariant sequences at the 5' end.

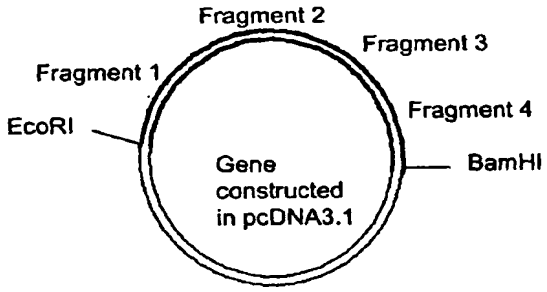


Paired 80mer oligos are connected via PCR in a stepwise manner from 5' to 3' using primers complimentary to the invariant seq.



108bp PCR fragments cloned into pGEM-T and sequenced. Clones with the proper sequence will be cut with 2 restriction enzymes. 4 fragments will be ligated together with pcDNA3.1 in a stepwise manner from the 5' to 3' end of gene

Fragments to be ligated with pcDNA3.1 (1-4 are in order from 5' to 3')	Restriction Enzymes Used to Cleave Fragment
Fragment 1	EcoRI/BsmBI
Fragment 2	BbsI/BsmBI
Fragment 3	BbsI/BsmBI
Fragment 4	BbsI/BamHI
pcDNA3.1	EcoRI/BamHI



Ligations will be repeated stepwise 5' to 3' until the entire gene has been cloned into pcDNA3.1

Figure 7

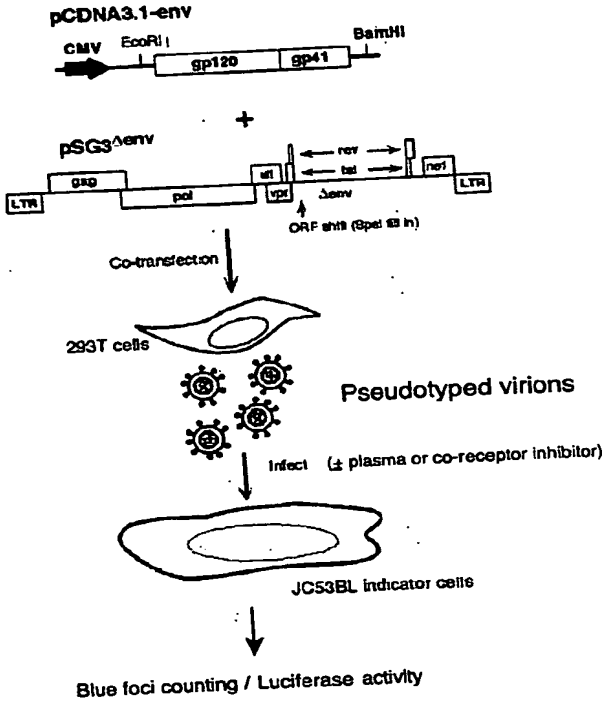
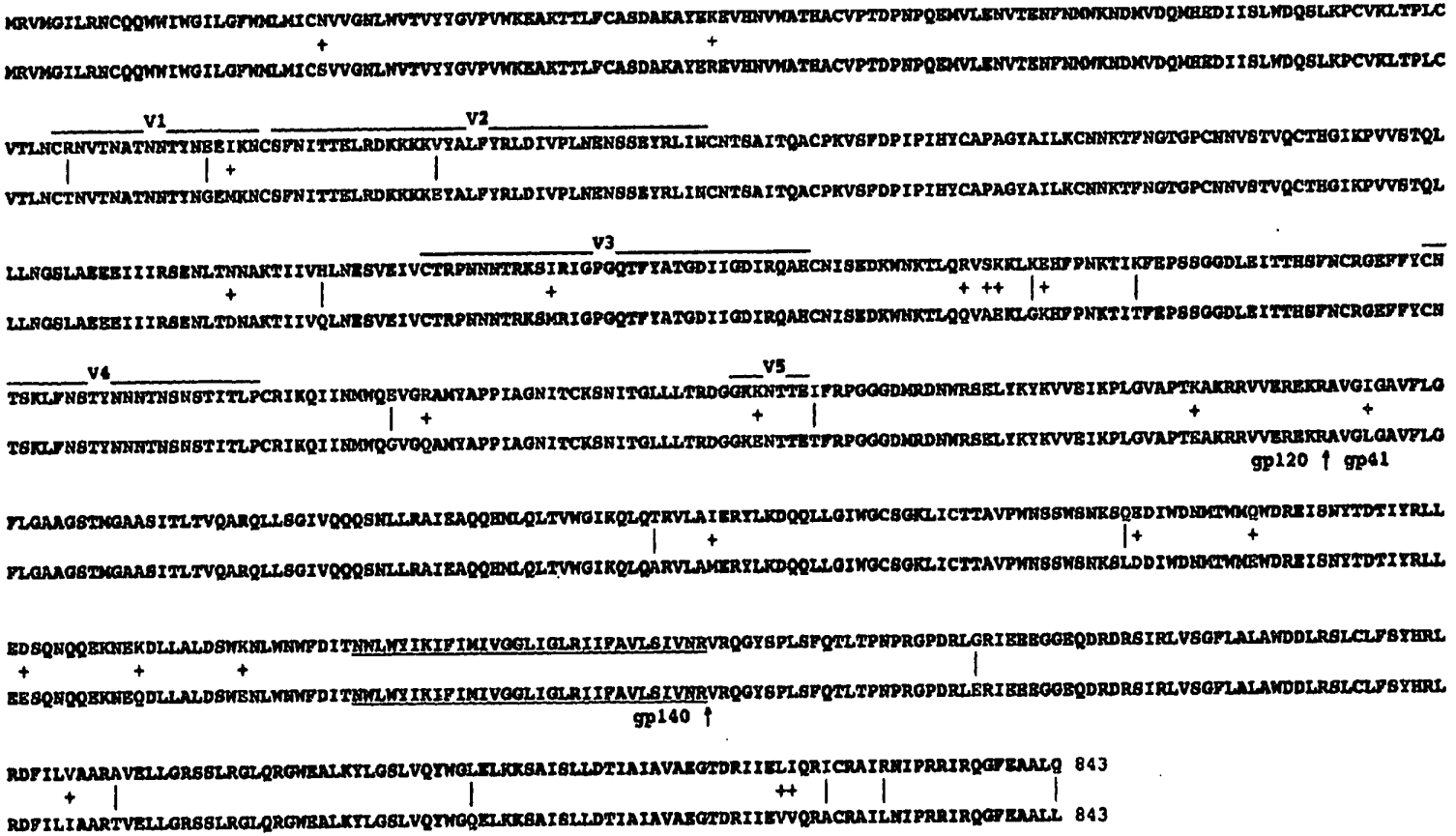

pCDNA3.1-env
CMV
EcoRI
gp120
gp41
BamHI
+
pSG3Δenv
LTR
gag
pol
vif
vpr
rev
tat
Δenv
nef
LTR
ORF shift (SpeI at in)
Co-transfection
293T cells
Pseudotyped virions
Infect (± plasma or co-receptor inhibitor)
JC53BL indicator cells
Blue foci counting / Luciferase activity

Figure 8



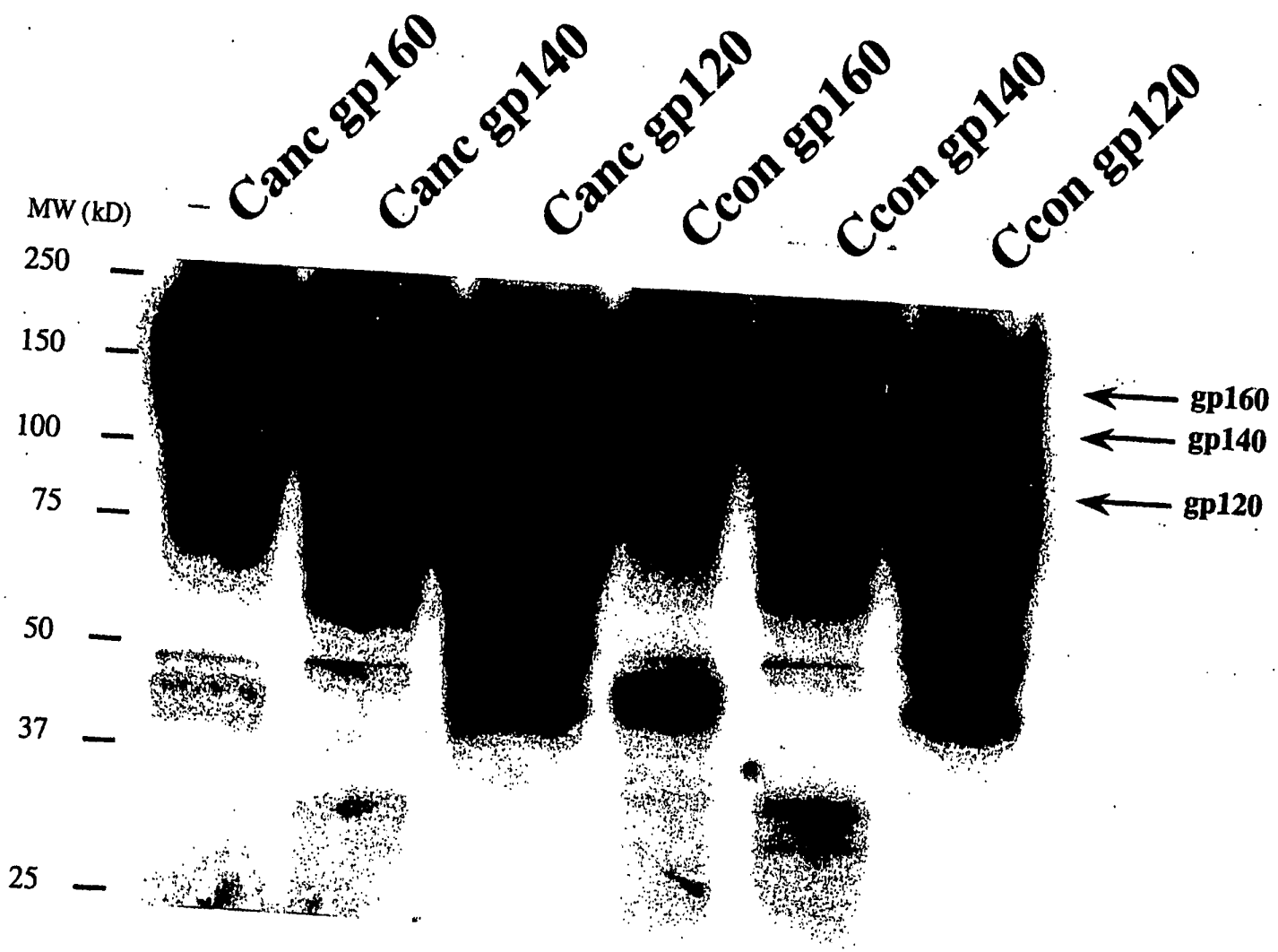

Canc gp160
Canc gp140
Canc gp120
Ccon gp160
Ccon gp140
Ccon gp120
MW (kD)
250
150
100
75
50
37
25
gp160
gp140
gp120

Figure 9

Figure 10 A

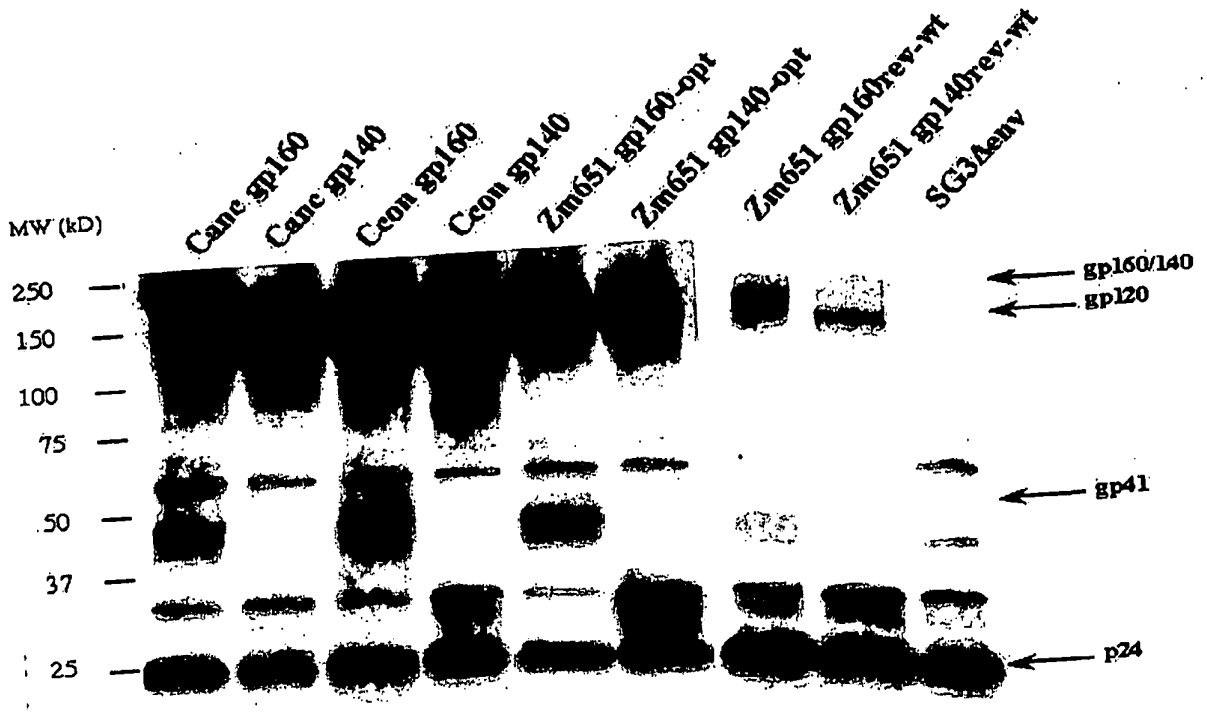

MW (kD)
Cane gp160
Cane gp140
Ceon gp160
Ceon gp140
Zm651 gp160-opt
Zm651 gp140-opt
Zm651 gp160rev-wt
Zm651 gp140rev-wt
SG3Δenv
250
150
100
75
50
37
25
gp160/140
gp120
gp41
p24

Figure 10B

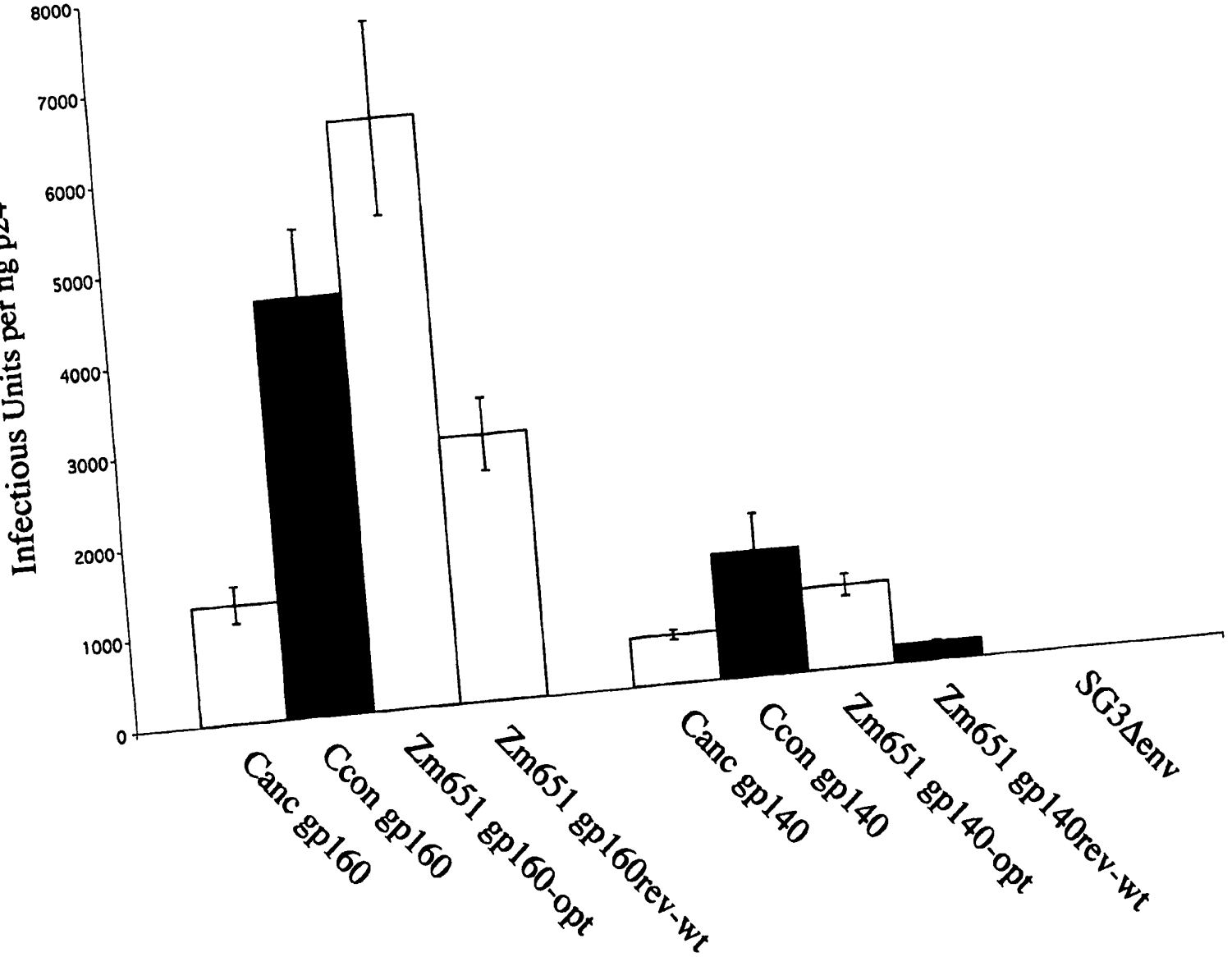

Infectious Units per ng p24
8000
7000
6000
5000
4000
3000
2000
1000
0
Canc gp160
Ccon gp160
Zm651 gp160-opt
Zm651 gp160rev-wt
Canc gp140
Ccon gp140
Zm651 gp140-opt
Zm651 gp140rev-wt
SG3Δenv

Figure 11

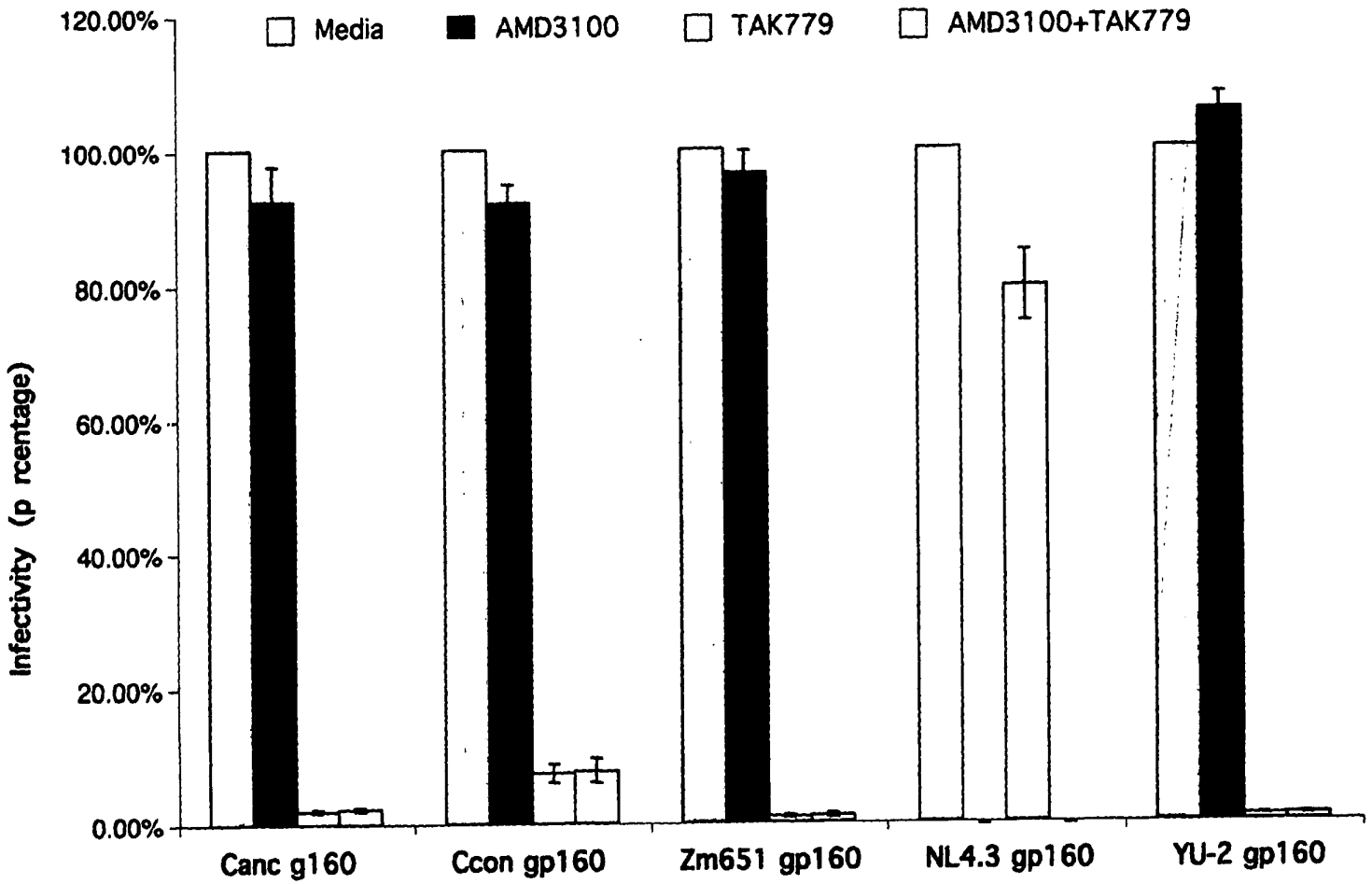

Media
AMD3100
TAK779
AMD3100+TAK779
120.00%
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
Infectivity (p rcentage)
Canc g160
Ccon gp160
Zm651 gp160
NL4.3 gp160
YU-2 gp160

Figure 12

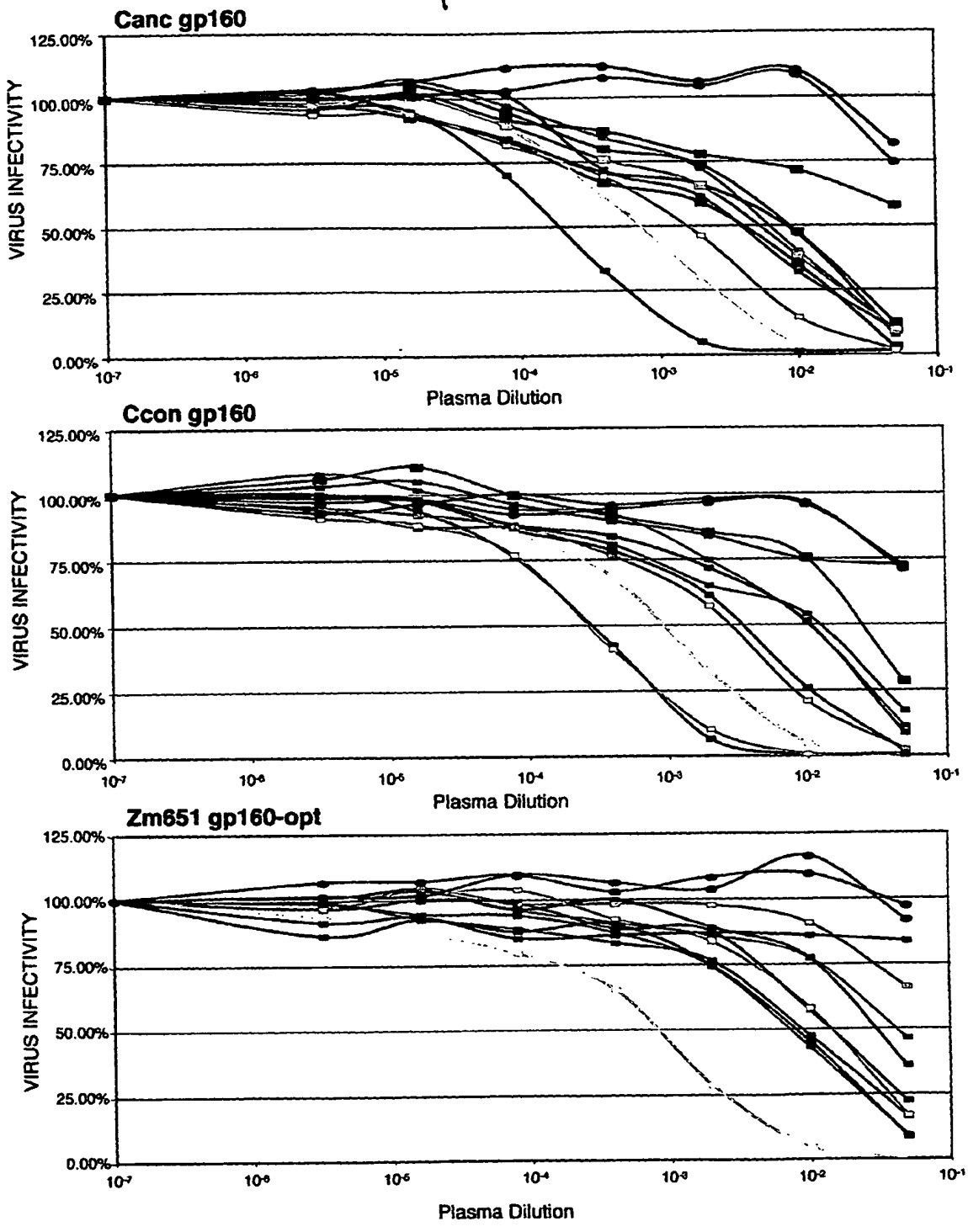
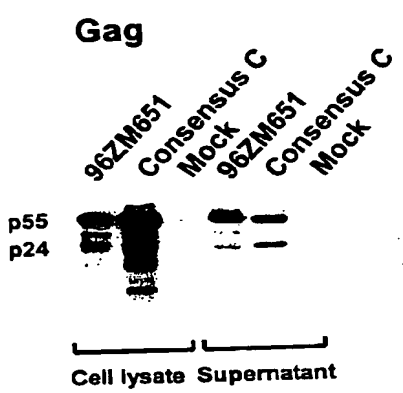

Canc gp160
VIRUS INFECTIVITY
125.00%
100.00%
75.00%
50.00%
25.00%
0.00%
10⁻⁷
10⁻⁶
10⁻⁵
10⁻⁴
10⁻³
10⁻²
10⁻¹
Plasma Dilution
A
Ccon gp160
VIRUS INFECTIVITY
Plasma Dilution
B
Zm651 gp160-opt
VIRUS INFECTIVITY
Plasma Dilution
C
Plasma from HIV-1 subtype C infected patients
Plasma from uninfected donors

Figure 13

A


Gag
96ZM651
Consensus C
Mock
96ZM651
Consensus C
Mock
p55
p24
Cell lysate
Supernatant

B

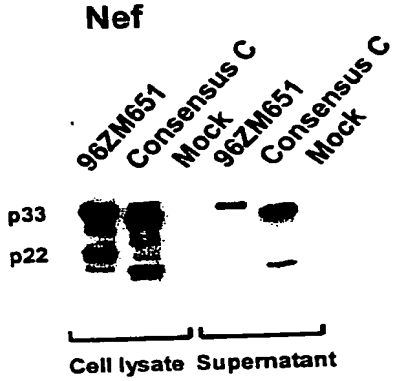

Nef
96ZM651
Consensus C
Mock
96ZM651
Consensus C
Mock
p33
p22
Cell lysate
Supernatant

Figure 13C

C.con.gag (subtype C consensus gag)
MGARASILRGGKLDTWEKIRLRPGGKKRYMIKHLVWASRELERFALNPGLLETSEGCKQIMKQLQPA
LQTGTEELRSLYNTVATLYCVHEKIEVRDTKEALDKIEEEQNKSQQKTQQAEAAADGKVSQNYPI
VQNLQGQMVHQAISPRTLNAWVKVIEEKAFSPEVIPMFTALSEGATPQDLNTMLNTVGGHQAAMQMLKDT
INEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIAWMTSNPPVPVGDIYKRWIILGLNKIV
RMYSPVSILDIKQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNANPDCKTILRALGPGASLE
EMMTACQGVGGPSHKARVLAEAMSQANNTNIMMQRSNFKGPKRIVKCFNCGKEGHIARNCRAPRKKGCWK
CGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPAESFRFEETTPA
PKQEPKDREPLTSLKSLFGSDPLSQ

C.con.nef (subtype C consensus nef)
MGGKWSKSSIVGWPAVRERIRRTEPAAEGVGAASQDLDKYGALTSSNTATNNADCAWLEAQEEEEEV
GFPVRPQVPLRPMTYKAAFDLSFFLKEKGGLEGLIYSKKRQEILDLWVYHTQGFFPDWQNYTPGPGVRYP
LTFGWCFKLVPVDPREVEEANEGENNCLLHPMSQHGMEDEDREVLKWKFDSHLARRHMARELHPEYYKDC

Figure 13D

Figure 13E

C.con.gag (subtype C consensus gag. Not in the public domain)
GCCGCCGCCATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACACCTGGGAGAAGATCCGCC
TGCGCCCCGGCGGCAAGAAGCGCTACATGATCAAGCACCTGGTGTGGGCCAGCCGCGAGCTGGAGCGCTT
CGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATCATGAAGCAGCTGCAGCCCGCC
CTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTACAACACCGTGGCCACCCTGTACTGCGTGCACGAGA
AGATCGAGGTGCGCGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAGAGCCAGCAGAA
GACCCAGCAGGCCGAGGCCGCCGCCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAG
GGCCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGG
CCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAGGACCTGAACAC
CATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCC
GCCGAGTGGGACCGCCTGCACCCCGTGCACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGCG
GCAGCGACATCGCCGGCACCACCAGCACCCTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCGT
GCCCGTGGGCGACATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACAGCCCC
GTGAGCATCCTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGA
CCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCCTGCTGGTGCAGAACGC
CAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCCCGGCGCCAGCCTGGAGGAGATGATGACCGCC
TGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTGCTGGCCGAGGCCATGAGCCAGGCCAACAACA
CCAACATCATGATGCAGCGCAGCAACTTCAAGGGCCCCAAGCGCATCGTGAAGTGCTTCAACTGCGGCAA
GGAGGGCCACATCGCCCGCAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGC
CACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCC
GCCCCGGCAACTTCCTGCAGAGCCGCCCCGAGCCCACCGCCCCCCCCGCCGAGAGCTTCCGCTTCGAGGA
GACCACCCCCGCCCCCAAGCAGGAGCCCAAGGACCGCGAGCCCCTGACCAGCCTGAAGAGCCTGTTCGGC
AGCGACCCCCTGAGCCAGTAA

C.con.nef (subtype C consensus nef. Not in the public domain)
GCCGCCGCCATGGGCGGCAAGTGGAGCAAGAGCAGCATCGTGGGCTGGCCCGCCGTGCGCGAGCGCATCC
GCCGCACCGAGCCCGCCGCCGAGGGCGTGGGCGCCGCCAGCCAGGACCTGGACAAGTACGGCGCCCTGAC
CAGCAGCAACACCGCCACCAACAACGCCGACTGCGCCTGGCTGGAGGCCCAGGAGGAGGAGGAGGAGGTG
GGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGCCGCCTTCGACCTGAGCTTCT
TCCTGAAGGAGAAGGGCGGCCTGGAGGGCCTGATCTACAGCAAGAAGCGCCAGGAGATCCTGGACCTGTG
GGTGTACCACACCCAGGGCTTCTTCCCCGACTGGCAGAACTACACCCCCGGCCCCGGCGTGCGCTACCCC
CTGACCTTCGGCTGGTGCTTCAAGCTGGTGCCCGTGGACCCCCGCGAGGTGGAGGAGGCCAACGAGGGCG
AGAACAACTGCCTGCTGCACCCCATGAGCCAGCACGGCATGGAGGACGAGGACCGCGAGGTGCTGAAGTG
GAAGTTCGACAGCCACCTGGCCCGCCGCCACATGGCCCGCGAGCTGCACCCCGAGTACTACAAGGACTGC
TGA

Figure 13F

Figure 14A

CONs.env (group M consensus env gene. This one contain the consensus sequence for variable regions in env gene)
MRVRGIQRNCQHLWRWGTLILGMLMICSAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNV
WATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNVNVTN
TTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNNNNSSNYRLINCNTSAITQACPKVSF
EPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENITNN
AKTIIVQLNESVEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLRE
HFNNKTIIFKPSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNNNNTNDTITLPCRIKQIINM
WQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVKIEPLG
VAPTKAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHL
LQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTTVPWNSSWSNKSQDEIWDNMTWMEWEREI
NNYTDIIYSLIEESQNQQEKNEQELLALDKWASLWNWFDITNWLWYIKIFIMIVGGLIGLRIVFAVLSIV
NRVRQGYSPLSFQTLIPNPRGPDRPEGIEEEGGEQDRDRSIRLVNGFLALAWDDLRSLCLFSYHRLRDFI
LIAARTVELLGRKGLRRGWEALKYLWNLLQYWGQELKNSAISLLDTTAIAVAEGTDRVIEVVQRACRAIL
NIPRRIRQGLERALL

FIGURE 14B

CONs.env (gorup M consensus env gene. This one contain the consensus sequence for variable regions in env gene. The identical amino acid sequences as in the public domain)

GCCGCCGCCATGCGCGTGCGCGGCATCCAGCGCAACTGCCAGCACCTGTG
GCGCTGGGGCACCCTGATCCTGGGCATGCTGATGATCTGCTCCGCCGCCG
AGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCC
AACACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCC
AGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAAC
AACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTC
CCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCA
CCAACGTGAACGTGACCAACACCACCAACAACACCGAGGAGAAGGGCGAG
ATCAAGAACTGCTCCTTCAACATCACCACCGAGATCCGCGACAAGAAGCA
GAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCGACGACA
ACAACAACAACTCCTCCAACTACCGCCTGATCAACTGCAACACCTCCGCC
ATCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCCACTA
CTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCA
ACGGCACCGGCCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGC
ATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGA
GGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACGCCAAGACCA
TCATCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCCAAC
AACAACACCCGCAAGTCCATCCGCATCGGCCCCGGCCAGGCCTTCTACGC
CACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCG
GCACCAAGTGGAACAAGACCCTGCAGCAGGTGGCCAAGAAGCTGCGCGAG
CACTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGCGACCT
GGAGATCACCACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCA
ACACCTCCGGCCTGTTCAACTCCACCTGGATCGGCAACGGCACCAAGAAC
AACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCAGAT
CATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATCG
AGGGCAAGATCACCTGCAAGTCCAACATCACCGGCCTGCTGCTGACCCGC
GACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCCGCCCCGGCGG
CGGCGACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTACAAGGTGG
TGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGCCGCGTG
GTGGAGCGCGAGAAGCGCGCCGTGGGCATCGGCGCCGTGTTCCTGGGCTT
CCTGGGCGCCGCCGGCTCCACCATGGGCGCCGCCTCCATCACCCTGACCG
TGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCCAACCTG
CTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGG
CATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGG
ACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATCTGCACC
ACCACCGTGCCCTGGAACTCCTCCTGGTCCAACAAGTCCCAGGACGAGAT
CTGGGACAACATGACCTGGATGGAGTGGGAGCGCGAGATCAACAACTACA
CCGACATCATCTACTCCCTGATCGAGGAGTCCCAGAACCAGCAGGAGAAG
AACGAGCAGGAGCTGCTGGCCCTGGACAAGTGGGCCTCCCTGTGGAACTG
GTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCG
TGGGCGGCCTGATCGGCCTGCGCATCGTGTTCGCCGTGCTGTCCATCGTG
AACCGCGTGCGCCAGGGCTACTCCCCCCTGTCCTTCCAGACCCTGATCCC
CAACCCCCGCGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGCG
AGCAGGACCGCGACCGCTCCATCCGCCTGGTGAACGGCTTCCTGGCCCTG
GCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCGCCTGCG
CGACTTCATCCTGATCGCCGCCCGCACCGTGGAGCTGCTGGGCCGCAAGG
GCCTGCGCCGCGGCTGGGAGGCCCTGAAGTACCTGTGGAACCTGCTGCAG
TACTGGGGCCAGGAGCTGAAGAACTCCGCCATCTCCCTGCTGGACACCAC
CGCCATCGCCGTGGCCGAGGGCACCGACCGCGTGATCGAGGTGGTGCAGC
GCGCCTGCCGCGCCATCCTGAACATCCCCCGCCGCATCCGCCAGGGCCTG
GAGCGCGCCCTGCTGTAA

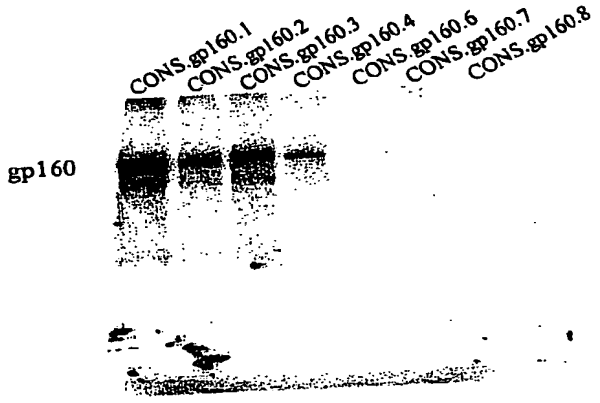
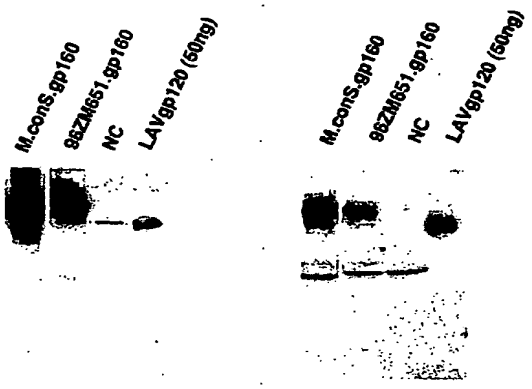

CONS.gp160.1
CONS.gp160.2
CONS.gp160.3
CONS.gp160.4
CONS.gp160.6
CONS.gp160.7
CONS.gp160.8
gp160

Figure 14C

Figure 15

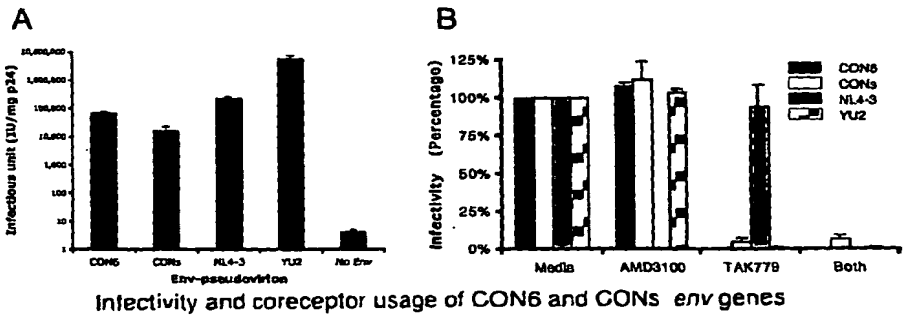
A



B

Cell lysate **Supernatant**

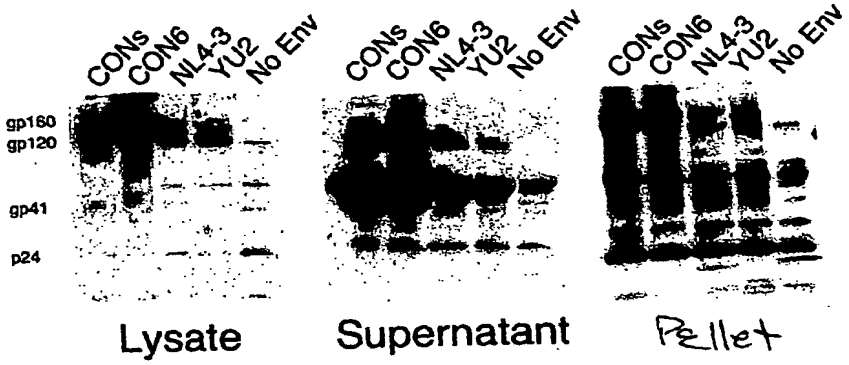
Expression of CONs *env* gene in mammalian cells



Infectivity and coreceptor usage of CON6 and CONs *env* genes

Figure 16

A B C



Env protein incorporation in CON6 and CONs Env-pseudovirions

Figure 17

Figure 18A

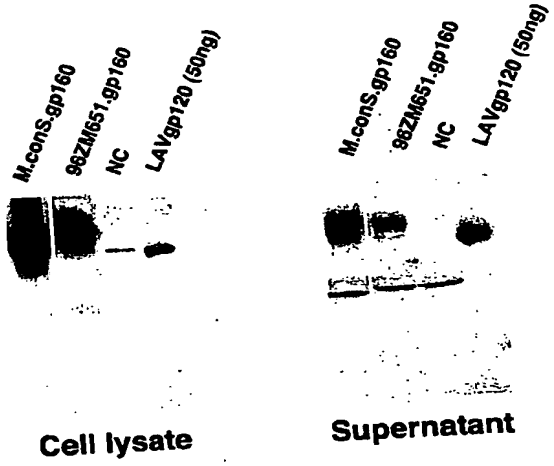
A.con.env (subtype A consensus env)
MRVMGIQRNCQHLWRWGTMILGMIIICSAAENLWVTVYYGVPVWKDAETTLFCASDAKAYDTEVHNV
WATHACVPTDPNPQEINLENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLNCSNVNVTT
NITNITDNMKGEIKNCSFNMTTELRDKKQKVYSLFYKLDVVQINKSNSSSQYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKEFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEVMIRSENITN
NAKNIIVQLTKPVKINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSRTEWNETLQKVAKQLR
KYFNNKTIIFTNSSGGDLEITTHSFNCGGEFFYCNTSGLFNSTWNGNGTKKKNSTESNDTITLPCRIKQI
INMWQRVGQAMYAPPIQGVIRCESNITGLLLTRDGGDNNSKNETFRPGGGDMRDNWRSELYKYKVVKIEP
LGVAPTKAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQ
HLLKLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTNVPWNSSWSNKSQSEIWDNMTWLQWDK
EISNYTDIIYNLIEESQNQQEKNEQDLLALDKWANLWNWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLS
VINRVRQGYSPLSFQTHTPNPGGLDRPGRIEEEGGEQGRDRSIRLVSGFLALAWDDLRSLCLFSYHRLRD
FILIAARTVELLGHSSLKGLRLGWEGLKYLWNLLLYWGRELKISAINLLDTIAIAVAGWTDRVIEIGQRI
CRAILNIPRRIRQGLERALL

Figure 18B

A.con.env (subtype A consensus env. Identical amino acid sequence to that in the public domain)

GCCGCCGCCATGCGCGTGATGGGCATCCAGCGCAACTGCCAGCACCTGTG
GCGCTGGGGCACCATGATCCTGGGCATGATCATCATCTGCTCCGCCGCCG
AGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGACGCC
GAGACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCC
AGGAGATCAACCTGGAGAACGTGACCGAGGAGTTCAACATGTGGAAGAAC
AACATGGTGGAGCAGATGCACACCGACATCATCTCCCTGTGGGACCAGTC
CCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCT
CCAACGTGAACGTGACCACCAACATCACCAACATCACCGACAACATGAAG
GGCGAGATCAAGAACTGCTCCTTCAACATGACCACCGAGCTGCGCGACAA
GAAGCAGAAGGTGTACTCCCTGTTCTACAAGCTGGACGTGGTGCAGATCA
ACAAGTCCAACTCCTCCTCCCAGTACCGCCTGATCAACTGCAACACCTCC
GCCATCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCCA
CTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAAGGACAAGGAGT
TCAACGGCACCGGCCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCAC
GGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGC
CGAGGAGGAGGTGATGATCCGCTCCGAGAACATCACCAACAACGCCAAGA
ACATCATCGTGCAGCTGACCAAGCCCGTGAAGATCAACTGCACCCGCCCC
AACAACAACACCCGCAAGTCCATCCGCATCGGCCCCGGCCAGGCCTTCTA
CGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACGTGT
CCCGCACCGAGTGGAACGAGACCCTGCAGAAGGTGGCCAAGCAGCTGCGC
AAGTACTTCAACAACAAGACCATCATCTTCACCAACTCCTCCGGCGGCGA
CCTGGAGATCACCACCCACTCCTTCAACTGCGGCGGCGAGTTCTTCTACT
GCAACACCTCCGGCCTGTTCAACTCCACCTGGAACGGCAACGGCACCAAG
AAGAAGAACTCCACCGAGTCCAACGACACCATCACCCTGCCCTGCCGCAT
CAAGCAGATCATCAACATGTGGCAGCGCGTGGGCCAGGCCATGTACGCCC
CCCCCATCCAGGGCGTGATCCGCTGCGAGTCCAACATCACCGGCCTGCTG
CTGACCCGCGACGGCGGCGACAACAACTCCAAGAACGAGACCTTCCGCCC
CGGCGGCGGCGACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTACA
AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGC
CGCGTGGTGGAGCGCGAGAAGCGCGCCGTGGGCATCGGCGCCGTGTTCCT
GGGCTTCCTGGGCGCCGCCGGCTCCACCATGGGCGCCGCCTCCATCACCC
TGACCGTGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCC
AACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGAAGCTGACCGT
GTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACC
TGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATC
TGCACCACCAACGTGCCCTGGAACTCCTCCTGGTCCAACAAGTCCCAGTC
CGAGATCTGGGACAACATGACCTGGCTGCAGTGGGACAAGGAGATCTCCA
ACTACACCGACATCATCTACAACCTGATCGAGGAGTCCCAGAACCAGCAG
GAGAAGAACGAGCAGGACCTGCTGGCCCTGGACAAGTGGGCCAACCTGTG
GAACTGGTTCGACATCTCCAACTGGCTGTGGTACATCAAGATCTTCATCA
TGATCGTGGGCGGCCTGATCGGCCTGCGCATCGTGTTCGCCGTGCTGTCC
GTGATCAACCGCGTGCGCCAGGGCTACTCCCCCCTGTCCTTCCAGACCCA
CACCCCCAACCCCGGCGGCCTGGACCGCCCCGGCCGCATCGAGGAGGAGG
GCGGCGAGCAGGGCCGCGACCGCTCCATCCGCCTGGTGTCCGGCTTCCTG
GCCCTGGCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCG
CCTGCGCGACTTCATCCTGATCGCCGCCCGCACCGTGGAGCTGCTGGGCC
ACTCCTCCCTGAAGGGCCTGCGCCTGGGCTGGGAGGGCCTGAAGTACCTG
TGGAACCTGCTGCTGTACTGGGGCCGCGAGCTGAAGATCTCCGCCATCAA
CCTGCTGGACACCATCGCCATCGCCGTGGCCGGCTGGACCGACCGCGTGA
TCGAGATCGGCCAGCGCATCTGCCGCGCCATCCTGAACATCCCCCGCCGC
ATCCGCCAGGGCCTGGAGCGCGCCCTGCTGTAA

C D



Expression of A.con *env* gene in mammalian cells

Figure 18

Figure 19A

M.con.gag (group M consensus gag. Identical amino acid sequence to that in the public domain)

GCCGCCGCCATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGA
CGCCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACCGCC
TGAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAAC
CCCGGCCTGCTGGAGACCTCCGAGGGCTGCAAGCAGATCATCGGCCAGCT
GCAGCCCGCCCTGCAGACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACA
CCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGAGGTGAAGGACACC
AAGGAGGCCCTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAA
GACCCAGCAGGCCGCCGCCGACAAGGGCAACTCCTCCAAGGTGTCCCAGA
ACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGGCCATC
TCCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCCTT
CTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCC
CCCAGGACCTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCC
ATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCG
CCTGCACCCCGTGCACGCCGGCCCCATCCCCCCCGGCCAGATGCGCGAGC
CCCGCGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGCAGATC
GCCTGGATGACCTCCAACCCCCCCATCCCCGTGGGCGAGATCTACAAGCG
CTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCCCCCGTGT
CCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTG
GACCGCTTCTTCAAGACCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAA
GAACTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCA
AGACCATCCTGAAGGCCCTGGGCCCCGGCGCCACCCTGGAGGAGATGATG
ACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTGCTGGC
CGAGGCCATGTCCCAGGTGACCAACGCCGCCATCATGATGCAGCGCGGCA
ACTTCAAGGGCCAGCGCCGCATCATCAAGTGCTTCAACTGCGGCAAGGAG
GGCCACATCGCCCGCAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAA
GTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCA
ACTTCCTGGGCAAGATCTGGCCCTCCAACAAGGGCCGCCCCGGCAACTTC
CTGCAGTCCCGCCCCGAGCCCACCGCCCCCCCCGCCGAGTCCTTCGGCTT
CGGCGAGGAGATCACCCCCTCCCCCAAGCAGGAGCCCAAGGACAAGGAGC
CCCCCCTGACCTCCCTGAAGTCCCTGTTCGGCAACGACCCCCTGTCCCAG
TAA

FIGURE 19B

M.con.pol.nuc
GCCGCCGCCatgccccagatcaccctgtggcagcgccccctggtgaccat
caagatcggcggccagctgaaggaggccctgctggccaccggcgccgacg
acaccgtgctggaggagatcaacctgcccggcaagtggaagcccaagatg
atcggcggcatcggcggcttcatcaaggtgcgccagtacgaccagatcct
gatcgagatctgcggcaagaaggccatcggcaccgtgctggtgggcccca
cccccgtgaacatcatcggccgcaacatgctgacccagatcggctgcacc
ctgaacttccccatctcccccatcgagaccgtgcccgtgaagctgaagcc
cggcatggacggccccaaggtgaagcagtggcccctgaccgaggagaaga
tcaaggccctgaccgagatctgcaccgagatggagaaggagggcaagatc
tccaagatcggccccgagaacccctacaacacccccatcttcgccatcaa
gaagaaggactccaccaagtggcgcaagctggtggacttccgcgagctga
acaagcgcacccaggacttctgggaggtgcagctgggcatcccccacccc
gccggcctgaagaagaagaagtccgtgaccgtgctggacgtgggcgacgc
ctacttctccgtgcccctggacgaggacttccgcaagtacaccgccttca
ccatcccctccatcaacaacgagacccccggcatccgctaccagtacaac
gtgctgccccagggctggaagggctcccccgccatcttccagtcctccat
gaccaagatcctggagcccttccgcacccagaaccccgagatcgtgatct
accagtacatggacgacctgtacgtgggctccgacctggagatcggccag
caccgcgccaagatcgaggagctgcgcgagcacctgctgcgctggggctt
caccaccccccgacaagaagcaccagaaggagcccccttcctgtggatgg
gctacgagctgcaccccgacaagtggaccgtgcagcccatccagctgccc
gagaaggactcctggaccgtgaacgacatccagaagctggtgggcaagct
gaactgggcctcccagatctaccccggcatcaaggtgaagcagctgtgca
agctgctgcgcggcgccaaggccctgaccgacatcgtgcccctgaccgag
gaggccgagctggagctggccgagaaccgcgagatcctgaaggagcccgt
gcacggcgtgtactacgacccctccaaggacctgatcgccgagatccaga
agcagggccaggaccagtggacctaccagatctaccaggagcccttcaag
aacctcaagaccggcaagtacgccaagatgcgctccgcccacaccaacga
cgtgaagcagctgaccgaggccgtgcagaagatcgccaccgagtccatcg
tgatctggggcaagaccccaagttccgcctgcccatccagaaggagacc

Figure 19B
continued

tgggagacctggtggaccgagtactggcaggccacctggattcccgagtg
ggagttcgtgaacacccccccctggtgaagctgtggtaccagctggaga
aggagcccatcgccggcgccgagaccttctacgtggacggcgccgccaac
cgcgagaccaagctgggcaaggccggctacgtgaccgaccgcggccgcca
gaaggtggtgtccctgaccgagaccaccaaccagaaaaccgagctgcagg
ccatccacctggccctgcaggactccggctccgaggtgaacatcgtgacc
gactcccagtacgccctgggcatcatccaggcccagcccgacaagtccga
gtccgagctggtgaaccagatcatcgagcagctgatcaagaaggagaagg
tgtacctgtcctgggtgcccgcccacaagggcatcggcggcaacgagcag
gtggacaagctggtgtccaccggcatccgcaaggtgctgttcctggacgg
catcgacaaggcccaggaggagcacgagaagtaccactccaactggcgcg
ccatggcctccgacttcaacctgccccccatcgtggccaaggagatcgtg
gcctcctgcgacaagtgccagctgaagggcgaggccatgcacggccaggt
ggactgctcccccggcatctggcagctggactgcacccacctggagggca
agatcatcctggtggccgtgcacgtggcctccggctacatcgaggccgag
gtgatccccgccgagaccggccaggagaccgcctacttcatcctgaagct
ggccggccgctggcccgtgaaggtgatccacaccgacaacggctccaact
tcacctccgccgccgtgaaggccgcctgctggtgggccggcatccagcag
gagttcggcatcccctacaaccccagtcccagggcgtggtggagtccat
gaacaaggagctgaagaagatcatcggccaggtgcgcgaccaggccgagc
acctcaagaccgccgtgcagatggccgtgttcatccacaacttcaagcgc
aagggcggcatcggcggctactccgccggcgagcgcatcatcgacatcat
cgccaccgacatccagaccaaggagctgcagaagcagatcaccaagatcc
agaacttccgcgtgtactaccgcgactcccgcgaccccatctggaagggc
cccgccaagctgctgtggaagggcgagggcgccgtggtgatccaggacaa
ctccgacatcaaggtggtgccccgccgcaaggccaagatcatccgcgact
acggcaagcagatggccggcgacgactgcgtggccggccgccaggacgag
gacTAA

Figure 19C

M.con.nef (group M consensus nef. Identical amino acid sequence to that in the public domain)
GCCGCCGCCATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCC
CGCCGTGCGCGAGCGCATCCGCCGCACCCACCCCGCCGCCGAGGGCGTGG
GCGCCGTGTCCCAGGACCTGGACAAGCACGGCGCCATCACCTCCTCCAAC
ACCGCCGCCAACAACCCCGACTGCGCCTGGCTGGAGGCCCAGGAGGAGGA
GGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGA
CCTACAAGGCCGCCCTGGACCTGTCCCACTTCCTGAAGGAGAAGGGCGGC
CTGGAGGGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTG
GGTGTACCACACCCAGGGCTACTTCCCCGACTGGCAGAACTACACCCCCG
GCCCCGGCATCCGCTACCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGTG
CCCGTGGACCCCGAGGAGGTGGAGGAGGCCAACGAGGGCGAGAACAACTC
CCTGCTGCACCCCATGTGCCAGCACGGCATGGAGGACGAGGAGCGCGAGG
TGCTGATGTGGAAGTTCGACTCCCGCCTGGCCCTGCGCCACATCGCCCGC
GAGCTGCACCCCGAGTACTACAAGGACTGCTAA

FIGURE 19D

C.con.pol.nuc

GCCGCCGCCatgccccagatcaccctgtggcagcgccccctggtgtccat
caaggtgggcggccagatcaaggaggccctgctggccaccggcgccgacg
acaccgtgctggaggagatcaacctgcccggcaagtggaagcccaagatg
atcggcggcatcggcggcttcatcaaggtgcgccagtacgaccagatcct
gatcgagatctgcggcaagaaggccatcggcaccgtgctggtgggcccca
cccccgtgaacatcatcggccgcaacatgctgacccagctgggctgcacc
ctgaacttccccatctcccccatcgagaccgtgcccgtgaagctgaagcc
cggcatggacggccccaaggtgaagcagtggcccctgaccgaggagaaga
tcaaggccctgaccgccatctgcgaggagatggagaaggagggcaagatc
accaagatcggccccgagaaccctacaacaccccgtgttcgccatcaa
gaagaaggactccaccaagtggcgcaagctggtggacttccgcgagctga
acaagcgcacccaggacttctgggaggtgcagctgggcatccccacccc
gccggcctgaagaagaagaagtccgtgaccgtgctggacgtgggcgacgc
ctacttctccgtgccctggacgagggcttccgcaagtacaccgccttca
ccatccctccatcaacaacgagaccccggcatccgctaccagtacaac
gtgctgccccagggctggaagggctcccccgccatcttccagtcctccat
gaccaagatcctggagcccttccgcgcccagaaccccgagatcgtgatct
accagtacatggacgacctgtacgtgggctccgacctggagatcggccag
caccgcgccaagatcgaggagctgcgcgagcacctgctgaagtggggctt
caccaccccgacaagaagcaccagaaggagccccccttcctgtggatgg
gctacgagctgcaccccgacaagtggaccgtgcagcccatccagctgccc
gagaaggactcctggaccgtgaacgacatccagaagctggtgggcaagct
gaactgggcctcccagatctaccccggcatcaaggtgcgccagctgtgca
agctgctgcgcggcgccaaggccctgaccgacatcgtgccccctgaccgag
gaggccgagctggagctggccgagaaccgcgagatcctgaaggagcccgt
gcacggcgtgtactacgaccccctccaaggacctgatcgccgagatccaga
agcagggccacgaccagtggacctaccagatctaccaggagcccttcaag
aacctcaagaccggcaagtacgccaagatgcgcaccgcccacaccaacga
cgtgaagcagctgaccgaggccgtgcagaagatcgccatggagtccatcg
tgatctggggcaagacccccaagttccgcctgcccatccagaaggagacc
tgggagacctggtggaccgactactggcaggccacctggattcccgagtg
ggagttcgtgaacaccccccccctggtgaagctgtggtaccagctggaga
aggagcccatcgccggcgccgagaccttctacgtggacggcgccgccaac
cgcgagaccaagatcggcaaggccggctacgtgaccgaccgcggccgcca
gaagatcgtgtccctgaccgagaccaccaaccagaaaaccgagctgcagg
ccatccagctggccctgcaggactccggctccgaggtgaacatcgtgacc
gactcccagtacgccctgggcatcatccaggcccagcccgacaagtccga
gtccgagctggtgaaccagatcatcgagcagctgatcaagaaggagcgcg
tgtacctgtcctgggtgcccgcccacaagggcatcggcggcaacgagcag
gtggacaagctggtgtcctccggcatccgcaaggtgctgttcctggacgg
catcgacaaggcccaggaggagcacgagaagtaccactccaactggcgcg
ccatggcctccgagttcaacctgcccccatcgtggccaaggagatcgtg
gcctcctgcgacaagtgccagctgaagggcgaggccatgcacggccaggt
ggactgctccccggcatctggcagctggactgcacccacctggagggca
agatcatcctggtggccgtgcacgtggcctccggctacatcgaggccgag

Figure 19D
continued

gtgatccccgccgagaccggccaggagaccgcctacttcatcctgaagct
ggccggccgctggcccgtgaaggtgatccacaccgacaacggctccaact
tcacctccgccgccgtgaaggccgcctgctggtgggccggcatccagcag
gagttcggcatcccctacaaccccccagtcccagggcgtggtggagtccat
gaacaaggagctgaagaagatcatcggccaggtgcgcgaccaggccgagc
acctcaagaccgccgtgcagatggccgtgttcatccacaacttcaagcgc
aagggcggcatcggcggctactccgccggcgagcgcatcatcgacatcat
cgccaccgacatccagaccaaggagctgcagaagcagatcatcaagatcc
agaacttccgcgtgtactaccgcgactcccgcgaccccatctggaagggc
cccgccaagctgctgtggaagggcgagggcgccgtggtgatccaggacaa
ctccgacatcaaggtggtgccccgccgcaaggccaagatcatcaaggact
acggcaagcagatggccggcgccgactgcgtggccggccgccaggacgag
gacTAA

Figure 19E

M.con.gag (group M consensus gag)
MGARASVLSGGKLDAWEKIRLRPGGKKKYRLKHLVWASRELERFALNPGLLETSEGCKQIIGQLQPA
LQTGSEELRSLYNTVATLYCVHQRIEVKDTKEALEKIEEEQNKSQQKTQQAAADKGNSSKVSQNYPIVQN
LQGQMVHQAISPRTLNAWVKVIEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKDTINE
EAAEWDRLHPVHAGPIPPGQMREPRGSDIAGTTSTLQEQIAWMTSNPPIPVGEIYKRWIILGLNKIVRMY
SPVSILDIRQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNANPDCKTILKALGPGATLEEMM
TACQGVGGPGHKARVLAEAMSQVTNAAIMMQRGNFKGQRRIIKCFNCGKEGHIARNCRAPRKKGCWKCGK
EGHQMKDCTERQANFLGKIWPSNKGRPGNFLQSRPEPTAPPAESFGFGEEITPSPKQEPKDKEPPLTSLK
SLFGNDPLSQ

Figure 19F

M.con.pol (group M consensus pol)
MPQITLWQRPLVTIKIGGQLKEALLaTGADDTVLEEINLPGKWKPKMIGGIGGFIKVRQYDQILIEICGK
KAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALTEICTE
MEKEGKISKIGPENPYNTPIFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLD
VGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRTQNPEIVI
YQYMDDLYVGSDLEIGQHRAKIEELREHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKD
SWTVNDIQKLVGKLNWASQIYPGIKVKQLCKLLRGAKALTDIVPLTEEAELELAENREILKEPVHGVYYD
PSKDLIAEIQKQGQDQWTYQIYQEPFKNLKTGKYAKMRSAHTNDVKQLTEAVQKIATESIVIWGKTPKFR
LPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIAGAETFYVDGAANRETKLGKAGYVTD
RGRQKVVSLTETTNQKTELQAIHLALQDSGSEVNIVTDSQYALGIIQAQPDKSESELVNQIIEQLIKKEK
VYLSWVPAHKGIGGNEQVDKLVSTGIRKVLFLDGIDKAQEEHEKYHSNWRAMASDFNLPPIVAKEIVASC
DKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPV
KVIHTDNGSNFTSAAVKAACWWAGIQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAV
FIHNFKRKGGIGGYSAGERIIDIIATDIQTKELQKQITKIQNFRVYYRDSRDPIWKGPAKLLWKGEGAVV
IQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED

Figure 19G

M.con.nef (group M consensus nef)
MGGKWSKSSIVGWPAVRERIRRTHPAAEGVGAVSQDLDKHGAITSSNTAANNPDCAWLEAQEEEEEVGFP
VRPQVPLRPMTYKAALDLSHFLKEKGGLEGLIYSKKRQEILDLWVYHTQGYFPDWQNYTPGPGIRYPLTF
GWCFKLVPVDPEEVEEANEGENNSLLHPMCQHGMEDEEREVLMWKFDSRLALRHIARELHPEYYKDC

Figure 19H

C.con.pol (subtype C consensus pol)

MPQITLWQRPLVSIKVGGQIKEALLaTGADDTVLEEINLPGKWKPKMIGGIGGFIKVRQYDQILIEICGK
KAIGTVLVGPTPVNIIGRNMLTQLGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALTAICEE
MEKEGKITKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLD
VGDAYFSVPLDEGFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRAQNPEIVI
YQYMDDLYVGSDLEIGQHRAKIEELREHLLKWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKD
SWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGAKALTDIVPLTEEAELELAENREILKEPVHGVYYD
PSKDLIAEIQKQGHDQWTYQIYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIAMESIVIWGKTPKFR
LPIQKETWETWWTDYWQATWIPEWEFVNTPPLVKLWYQLEKEPIAGAETFYVDGAANRETKIGKAGYVTD
RGRQKIVSLTETTNQKTELQAIQLALQDSGSEVNIVTDSQYALGIIQAQPDKSESELVNQIIEQLIKKER
VYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKAQEEHEKYHSNWRAMASEFNLPPIVAKEIVASC
DKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPV

KVIHTDNGSNFTSAAVKAACWWAGIQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAV
FIHNFKRKGGIGGYSAGERIIDIIATDIQTKELQKQIIKIQNFRVYYRDSRDPIWKGPAKLLWKGEGAVV
IQDNSDIKVVPRRKAKIIKDYGKQMAGADCVAGRQDED

Figure 20A

B.con.gag (subtype B consensus gag. The amino acid sequence is different from Los Alamos Database August 2002)

GCCGCCGCCATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCGAGCTGGA
CCGCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACAAGC
TGAAGCACATCGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCGTGAAC
CCCGGCCTGCTGGAGACCTCCGAGGGCTGCCGCCAGATCCTGGGCCAGCT
GCAGCCCTCCCTGCAGACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACA
CCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGAGGTGAAGGACACC
AAGGAGGCCCTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAA
GGCCCAGCAGGCCGCCGCCGACACCGGCAACTCCTCCCAGGTGTCCCAGA
ACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGGCCATC
TCCCCCCGCACCCTGAACGCCTGGGTGAAGGTGGTGGAGGAGAAGGCCTT
CTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCC
CCCAGGACCTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCC
ATGCAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCCGAGTGGGACCG
CCTGCACCCCGTGCACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGC
CCCGCGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGCAGATC
GGCTGGATGACCAACAACCCCCCCATCCCCGTGGGCGAGATCTACAAGCG
CTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCCCCCACCT
CCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTG
GACCGCTTCTACAAGACCCTGCGCGCCGAGCAGGCCTCCCAGGAGGTGAA
GAACTGGATGACCGAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCA
AGACCATCCTGAAGGCCCTGGGCCCCGCCGCCACCCTGGAGGAGATGATG
ACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTGCTGGC
CGAGGCCATGTCCCAGGTGACCAACTCCGCCACCATCATGATGCAGCGCG
GCAACTTCCGCAACCAGCGCAAGACCGTGAAGTGCTTCAACTGCGGCAAG
GAGGGCCACATCGCCAAGAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTG
GAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGG
CCAACTTCCTGGGCAAGATCTGGCCCTCCCACAAGGGCCGCCCCGGCAAC
TTCCTGCAGTCCCGCCCCGAGCCCACCGCCCCCCCCGAGGAGTCCTTCCG
CTTCGGCGAGGAGACCACCACCCCCTCCCAGAAGCAGGAGCCCATCGACA
AGGAGCTGTACCCCCTGGCCTCCCTGCGCTCCCTGTTCGGCAACGACCCC
TCCTCCCAGTAA

Figure 20 B

B.con.env (subtype B consensus env. The amino acid sequence is different from Los Alamos Database August 2002)

GCCGCCGCCATGCGCGTGAAGGGCATCCGCAAGAACTACCAGCACCTGTG
GCGCTGGGGCACCATGCTGCTGGGCATGCTGATGATCTGCTCCGCCGCCG
AGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCC
ACCACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCC
AGGAGGTGGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAAC
AACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTC
CCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCA
CCGACCTGAAGAACAACCTGCTGAACACCAACTCCTCCTCCGGCGAGAAG
ATGGAGAAGGGCGAGATCAAGAACTGCTCCTTCAACATCACCACCTCCAT
CCGCGACAAGGTGCAGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGG
TGCCCATCGACAACAACAACAACACCTCCTACCGCCTGATCTCCTGCAAC
ACCTCCGTGATCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCC
CATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACA
AGAAGTTCAACGGCACCGGCCCCTGCACCAACGTGTCCACCGTGCAGTGC
ACCCACGGCATCCGCCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTC
CCTGGCCGAGGAGGAGGTGGTGATCCGCTCCGAGAACTTCACCGACAACG
CCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGCACC
CGCCCCAACAACAACACCCGCAAGTCCATCCACATCGGCCCCGGCCGCGC
CTTCTACACCACCGGCGAGATCATCGGCGACATCCGCCAGGCCCACTGCA
ACATCTCCCGCGCCAAGTGGAACAACACCCTGAAGCAGATCGTGAAGAAG
CTGCGCGAGCAGTTCGGCAACAAGACCATCGTGTTCAACCAGTCCTCCGG
CGGCGACCCCGAGATCGTGATGCACTCCTTCAACTGCGGCGGCGAGTTCT
TCTACTGCAACACCACCCAGCTGTTCAACTCCACCTGGAACGACAACGGC
ACCTGGAACAACACCAAGGACAAGAACACCATCACCCTGCCCTGCCGCAT
CAAGCAGATCATCAACATGTGGCAGGAGGTGGGCAAGGCCATGTACGCCC
CCCCCATCCGCGGCCAGATCCGCTGCTCCTCCAACATCACCGGCCTGCTG
CTGACCCGCGACGGCGGCAACAACAACAACGACACCGAGATCTTCCGCCC
CGGCGGCGGCGACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTACA
AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGC
CGCGTGGTGCAGCGCGAGAAGCGCGCCGTGGGCATCGGCGCCATGTTCCT
GGGCTTCCTGGGCGCCGCCGGCTCCACCATGGGCGCCGCCTCCATGACCC
TGACCGTGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGAAC
AACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGT
GTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACC
TGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATC
TGCACCACCACCGTGCCCTGGAACGCCTCCTGGTCCAACAAGTCCCTGGA
CGAGATCTGGGACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACA
ACTACACCTCCCTGATCTACACCCTGATCGAGGAGTCCCAGAACCAGCAG
GAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCTCCCTGTG
GAACTGGTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCATCA
TGATCGTGGGCGGCCTGATCGGCCTGCGCATCGTGTTCGCCGTGCTGTCC
ATCGTGAACCGCGTGCGCCAGGGCTACTCCCCCCTGTCCTTCCAGACCCG
CCTGCCCGCCCCCCGCGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGG
GCGGCGAGCGCGACCGCGACCGCTCCGGCCGCCTGGTGGACGGCTTCCTG
GCCCTGATCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCG
CCTGCGCGACCTGCTGCTGATCGTGACCCGCATCGTGGAGCTGCTGGGCC
GCCGCGGCTGGGAGGTGCTGAAGTACTGGTGGAACCTGCTGCAGTACTGG
TCCCAGGAGCTGAAGAACTCCGCCGTGTCCCTGCTGAACGCCACCGCCAT
CGCCGTGGCCGAGGGCACCGACCGCGTGATCGAGGTGGTGCAGCGCGCCT
GCCGCGCCATCCTGCACATCCCCCGCCGCATCCGCCAGGGCCTGGAGCGC
GCCCTGCTGTAA

Figure 20C

B.con.gag (subtype B consensus gag)
MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQT
GSEELRSLYNTVATLYCVHQRIEVKDTKEALEKIEEEQNKSKKKAQQAAADTGNSSQVSQNYPIVQNLQG
QMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAA
EWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRMYSPT
SILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTAC
QGVGGPGHKARVLAEAMSQVTNSATIMMQRGNFRNQRKTVKCFNCGKEGHIAKNCRAPRKKGCWKCGKEG
HQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPEESFRFGEETTTPSQKQEPIDKELYPLASLR
SLFGNDPSSQ

B.con.env (subtype B consensus env)
MRVKGIRKNYQHLWRWGTMLLGMLMICSAAEKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWAT
HACVPTDPNPQEVVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLKNNLLNT
NSSSGEKMEKGEIKNCSFNITTSIRDKVQKEYALFYKLDVVPIDNNNNTSYRLISCNTSVITQACPKVSF
EPIPIHYCAPAGFAILKCNDKKFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSENFTDN
AKTIIVQLNESVEINCTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNNTLKQIVKKLRE
QFGNKTIVFNQSSGGDPEIVMHSFNCGGEFFYCNTTQLFNSTWNDNGTWNNTKDKNTITLPCRIKQIINM
WQEVGKAMYAPPIRGQIRCSSNITGLLLTRDGGNNNNDTEIFRPGGGDMRDNWRSELYKYKVVKIEPLGV
APTKAKRRVVQREKRAVGIGAMFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIEAQQHLL
QLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTTVPWNASWSNKSLDEIWDNMTWMEWEREID
NYTSLIYTLIEESQNQQEKNEQELLELDKWASLWNWFDITNWLWYIKIFIMIVGGLIGLRIVFAVLSIVN
RVRQGYSPLSFQTRLPAPRGPDRPEGIEEEGGERDRDRSGRLVDGFLALIWDDLRSLCLFSYHRLRDLLL
IVTRIVELLGRRGWEVLKYWWNLLQYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQRACRAILHIPRR
IRQGLERALL

Figure 20D

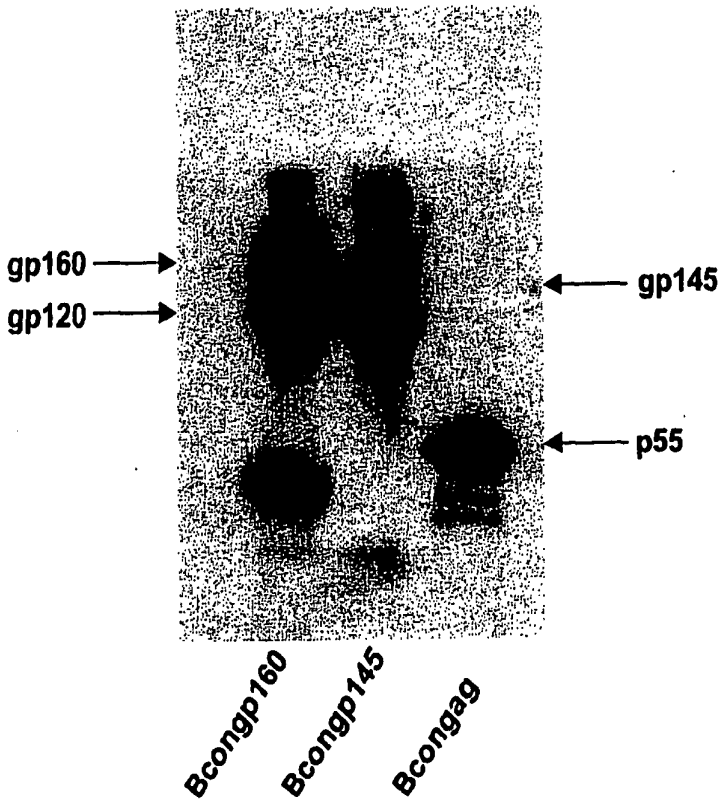
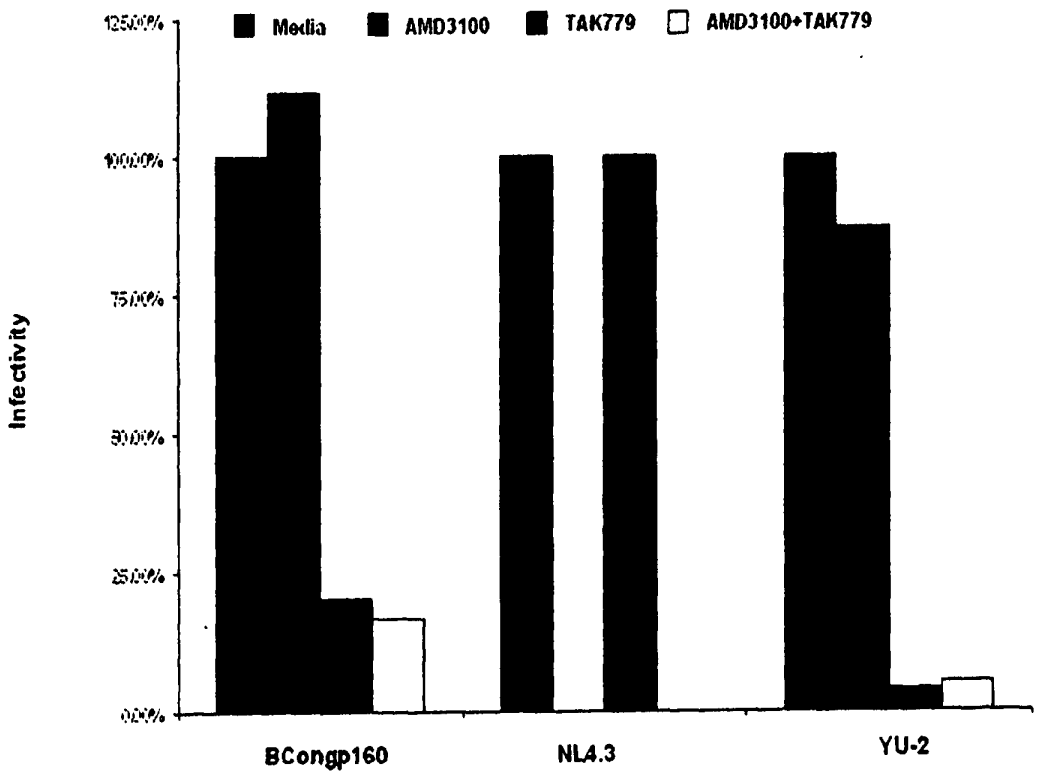


Figure 21. Expression of subtype B consensus *env* and *gag* genes in 293T cells. Plasmids containing codon-optimized subtype B consensus *gp160, gp140*, and *gag* genes were transfected into 293T cells, and protein expression was examined by Western Blot analysis of cell lysates. 48-hours post-transfection, cell lysates were collected, total protein content determined by the BCA protein assay, and 2 μg of total protein was loaded per lane on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with serum from an HIV-1 subtype B infected individual.

Figure 22



Co-receptor usage of subtype B consensus envelopes. Pseudotyped particles containing the subtype B consensus gp160 Env were incubated with DEAE-Dextran treated JC53-BL cells in the presence of AMD3100 (a specific inhibitor of CXCR4), TAK779 (a specific inhibitor of CCR5), and AMD3000+TAK779 to determine co-receptor usage. NL4.3, an isolate known to utilize CXCR4 and YU-2, a known CCR5-using isolate, were included as controls.

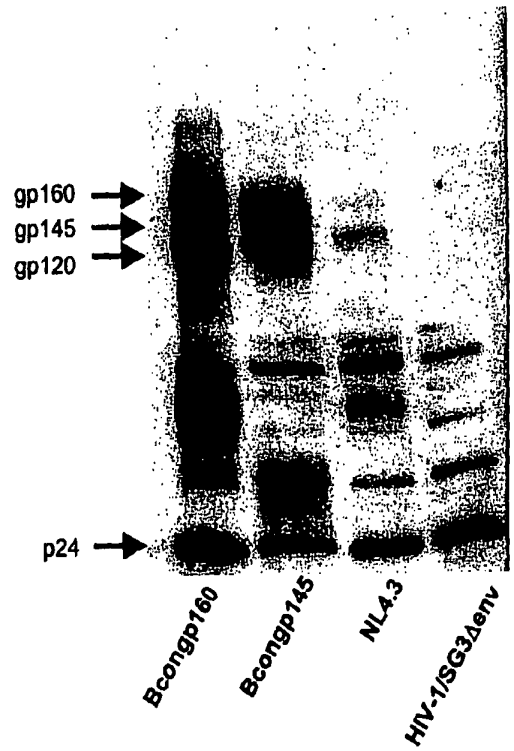


Figure 234. ***Trans* complementation of *env*-deficient HIV-1 with codon-optimized subtype B consensus *gp160* and *gp140* genes.**

Plasmids containing codon-optimized, subtype B consensus *gp160* or *gp140* genes were co-transfected into 293T cells with an HIV-1/SG3Δenv provirus. 48-hours post-transfection cell supernatants containing pseudotyped virus were harvested, clarified in a tabletop centrifuge, filtered through a 0.2μM filter, and pellet through a 20% sucrose cushion. Quantification of p24 in each virus pellet was determined using the Coulter HIV-1 p24 antigen assay; 25 ng of p24 was loaded per lane on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with anti-HIV-1 antibodies from infected HIV-1 subtype B patient serum. *Trans* complementation with a rev-dependent NL4.3 *env* was included for control.

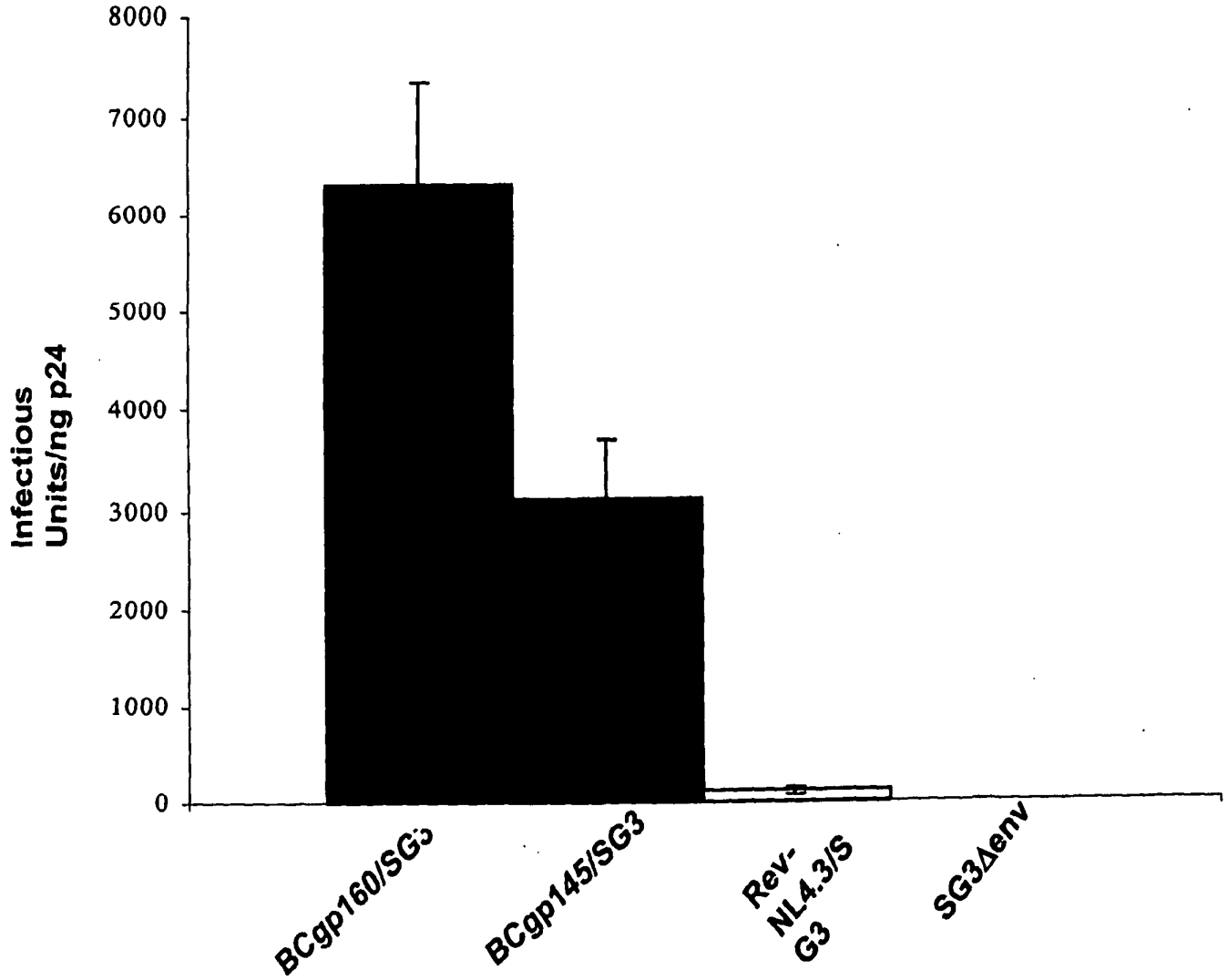
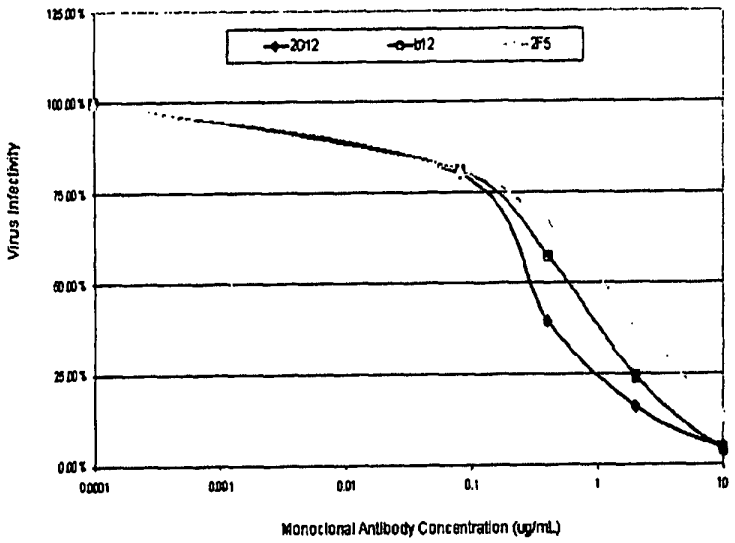


Figure 23B **Infectivity of virus particles containing the subtype B concensus envelope.**

Infectivity of pseudotyped virus containing consensus B gp160 or gp140 was determined using the JC53-BL assay. Sucrose cushion purified virus particles were assayed by the Coulter p24 antigen assay, and 5-fold serial dilutions of each pellet were incubated with DEAE-Dextran treated JC53-BL cells. Following a 48-hour incubation period, cells were fixed and stained to visualize β-galactosidase expressing cells. Infectivity is expressed as infectious units per ng of p24.

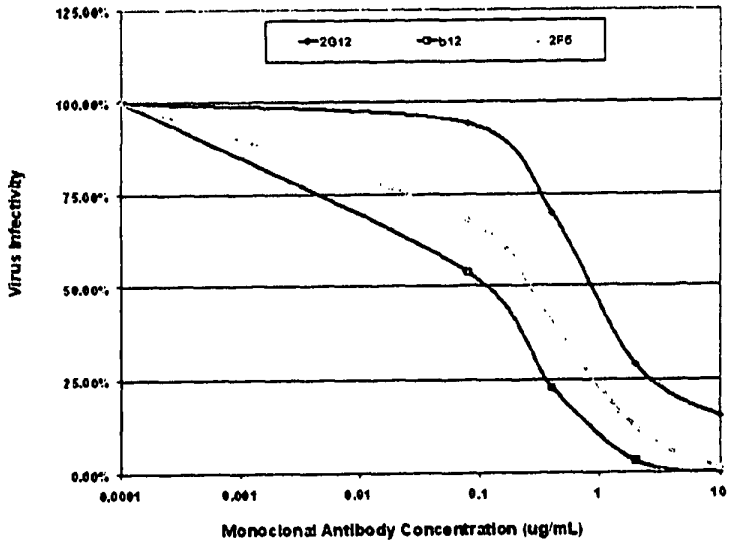
Figure 24

A



Neutralization of Pseudovirions containing Subtype B consensus Env (gp160)

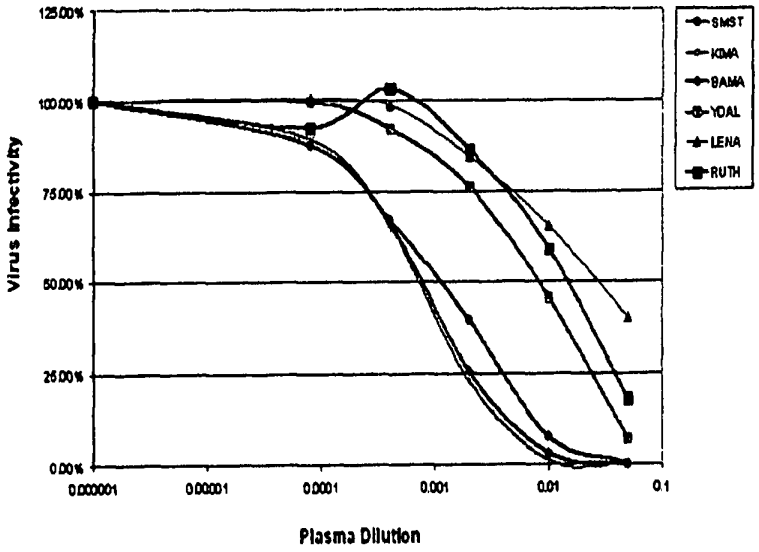
B



Neutralization of Pseudovirions containing NL4.3 Env (gp160)

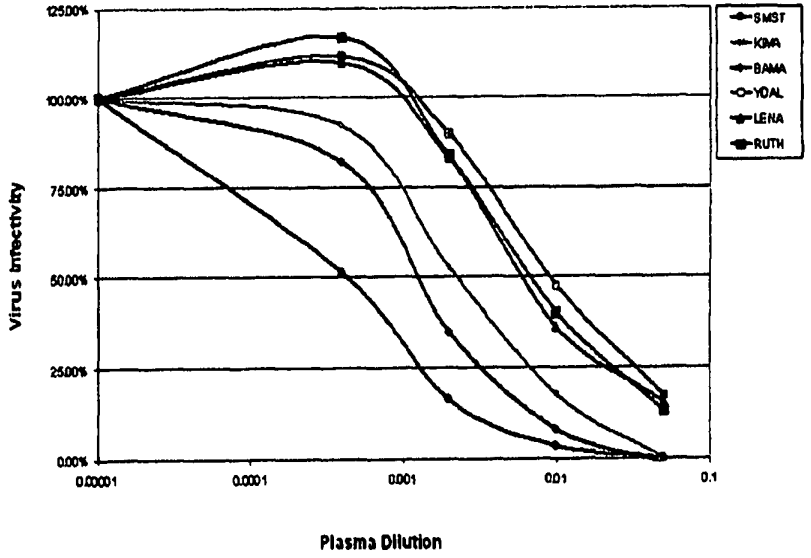
Figure 24

C



Neutralization of Pseudovirions containing Subtype B consensus Env (gp160)

D

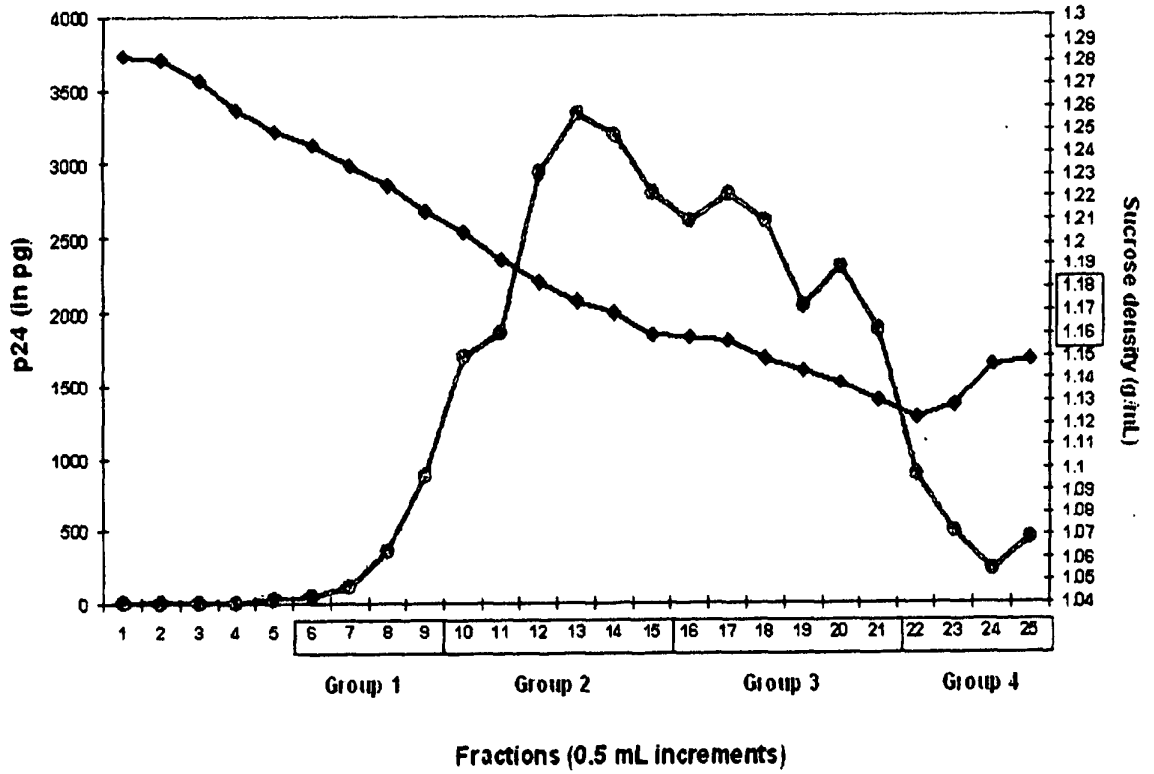


Neutralization of Pseudovirions containing NL4.3 Env (gp160)

Neutralization sensitivity of virions containing subtype B concensus gp 160 envelop .

Equivalent amounts of pseudovirions containing the subtype B consensus or NL4.3 Env (gp160) (1,500 infectious units) were preincubated with three different monoclonal neutralizing antibodies and a panel of plasma samples from HIV-1 subtype B infected individuals, and then added to the JC53-BL cell monolayer in 96-well plates. Plates were cultured for two days and luciferase activity was measured as an indicator of viral infectivity. Virus infectivity was calculated by dividing the luciferase units (LU) produced at each concentration of antibody by the LU produced by the control infection. The mean 50% inhibitory concentration (IC_{50}) and the actual % neutralization at each antibody dilution were then calculated for each virus. The results of all luciferase experiments were confirmed by direct counting of blue foci in parallel infections.

Figure 25 A



Density and p24 analysis of sucrose gradient fractions.

0.5ml fractions were collected from a 20-60% sucrose gradient. Fraction number 1 represents the most dense fraction taken from the bottom of the gradient tube. Density was measured with a refractometer and the amount of p24 in each fraction was determined by the Coulter p24 antigen assay. Fractions 6-9, 10-15, 16-21, and 22-25 were pooled together and analyzed by Western Blot. As expected, virions sedimented at a density of 1.16-1.18 g/ml.

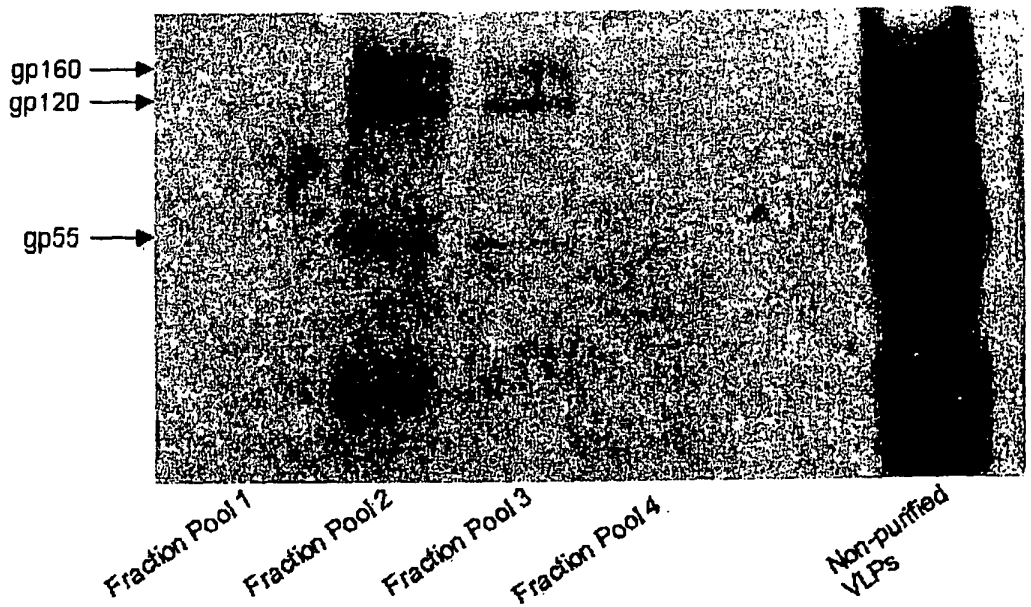


Figure 25B **VLP production by co-transfection of subtype B consensus *gag* and *env* genes.**

293T cells were co-transfected with subtype B consensus *gag* and *env* genes. Cell supernatants were harvested 48-hours post-transfection, clarified through at 20% sucrose cushion, and further purified through a 20-60% sucrose gradient. Select fractions from the gradient were pooled, added to 20ml of PBS, and centrifuged overnight at 100,000 x g. Resuspended pellets were loaded onto a 4-20% SDS-PAGE gel, proteins were transferred to a PVDF membrane, and probed with plasma from an HIV-1 subtype B infected individual.

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US2004/030397

International filing date: 17 September 2004 (17.09.2004)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/503,460
Filing date: 17 September 2003 (17.09.2003)

Date of receipt at the International Bureau: 10 May 2006 (10.05.2006)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse